SCIENCE LIBRARY

Vol. XVII, No. 3 *PSYCHOLOGICAL REVIEW PUBLICATIONS* JUN 1934 JUNE, 1934

Journal of Experimental Psychology

18

EDITED BY
SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

WALTER S. HUNTER, CLARK UNIVERSITY (*Psychol. Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Psychol. Rev.*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Psychol. Bull.*)
JOSEPH PETERSON, GEORGE PEABODY COLLEGE (*Psychol. Monog.*)

CONTENTS

An Experimental Attempt to Produce Artificial Chromaesthesia by the Technique of the Conditioned Response: E. L. KELLY 315

The Affective Value of Color as a Function of Hue, Tint and Chroma: J. P. GUILFORD 342

The Interrelations of Speed of Reaction Measurements: L. H. LANIER 371

Prior Entry Within a Single Sense Department: J. G. NEEDHAM 400

Further Experiments in the Strength of Connections Where the Connection is Punished or Rewarded or Neither Punished Nor Rewarded: I. LORGE, J. EISENSON AND B. EPSTEIN 412

The Rôle of Kinesthesis in Ideational Maze Learning: W. P. CHASE 424

Motivation in Learning. II. The Function of Electric Shock for Right and Wrong Responses in Human Subjects: K. F. MUENZINGER 439

A View of the Cortical Integrational Process Through Liminal Visual Stimuli: M. LUCKIESH AND F. K. MOSS 449

The Nature of Distance Vision: O. O. NORRIS 462

A Correlation Form with Absolute Checks: R. C. TRYON 477

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, University of Pennsylvania (*J. Exper. Psych.*)
WALTER S. HUNTER, Clark University (*Index and Abstracts*)
HENRY T. MOORE, Skidmore College (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, Princeton University (*Review*)
EDWARD S. ROBINSON, Yale University (*Bulletin*)
JOSEPH PETERSON, George Peabody College (*Monographs*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the Bulletin consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The Index is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The Monographs appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: \$5.50 (Foreign, \$5.75). **Index:** \$4.00 per volume.
Journal: \$7.00 (Foreign, \$7.25). **Monographs:** \$6.00 per volume (Foreign, \$6.30).
Bulletin: \$6.00 (Foreign, \$6.25). **Abstracts:** \$6.00 (Foreign, \$6.25).
Abnormal and Social: \$5.00 (Foreign, \$5.25). Single copies \$1.50.
Current numbers: **Journal,** \$1.25; **Review,** \$1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES

Review and Bulletin: \$10.00 (Foreign, \$10.50).
Review and J. Exp.: \$11.00 (Foreign, \$11.50).
Bulletin and J. Exp.: \$12.00 (Foreign, \$12.50).
Review, Bulletin, and J. Exp.: \$16.00 (Foreign, \$16.75).
Review, Bulletin, J. Exp., and Index: \$19.00 (Foreign, \$19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XVII, No. 3 JUNE, 1934

AN EXPERIMENTAL ATTEMPT TO PRODUCE ARTIFICIAL CHROMÆSTHESIA BY THE TECHNIQUE OF THE CONDITIONED RESPONSE [1]

BY E. LOWELL KELLY

Connecticut State College

I. INTRODUCTION

It is a well known although curious fact that some persons possess a tendency for perceiving certain types of stimuli, simultaneously, in more than one sense field. For some individuals, tones are not only heard but seen either as colors or as forms. Other persons find that letters and numbers, whether printed or spoken, have specific colors. Such a tendency toward double sensations is known as synæsthesia. While there is thus, theoretically, a large number of possible combinations of the various sense fields, it so happens that vision and audition are the sense modalities most frequently involved. The particular type of synæsthesia in which colors accompany auditory stimuli has been variously referred to as: *Farbenhören, audition colorée, Akustish-chromæsthesia Synopsie*, colored hearing, pseudo-chromæsthesia, and chromæsthesia. The current practice in the English literature, however, seems to favor the last mentioned term. Throughout this report,

[1] This experiment was conducted at Stanford University under the direction of Prof. Walter R. Miles.

the word 'chromæsthesia' will be understood to denote that tendency of auditory stimuli to arouse simultaneous sensations or images of color as well as of sound.

Although the literature bearing directly or indirectly on the subject of chromæsthesia is very extensive it consists largely of such contributions as: introspective reports, hearsay evidence, portrayals of individual cases, and similar anecdotal and descriptive articles. Inasmuch as several excellent summaries have been made of this literature, Krohn (**12**), Mahling (**16**), Argelander (**2**), and Wheeler (**27**), and since it has no direct bearing on experimental investigation in the field, it will be given but brief review.

One of the first accounts of this interesting phenomenon is to be found in Goethe's (**9**) *Theory of Colors*, written in 1810. Goethe refers to a case reported by Hoffman (**10**) as early as 1786. Hoffman describes the case of a Swiss magistrate and painter who was talented with the ability to see colors as accompaniments of sounds. For this individual, tones of musical instruments called forth unusually strong color sensations, the high notes resulting in especially vivid color impressions.

According to Krohn (**12**), the first case of chromæsthesia to be reported in the medical literature is the detailed description given by George Saches in his inaugural dissertation at Erlangen in 1812. The person he describes was an albino who associated color, not only with musical tones, but also with letters of the alphabet, figures, names of cities, days of the week, dates, epochs of history, and even with phases of human life. Of especial interest is the fact that, when a musical tone was associated with a color, the letter indicating that note was associated with the same color.

In 1843, Gautier (**8**) published additional descriptions of persons with tendencies toward chromæsthesia. Of greater significance was his report of being able to produce artificially these false color sensations by the use of the drug, hasheesh. This is particularly interesting since Havelock Ellis (**5**) records a similar tendency to artificial synæsthesia as one of the effects of the drug, mescal or peyote.

Probably the most interesting case study of chromæsthesia reported is that of Nussbaumer. When a student at Vienna, at the age of 23, he published (**18**) a description of both his own and his brother's chromæsthetic tendencies. A summary of this interesting self-description has been made by Lewes (**13**). Galton (**7**) in his studies of mental imagery discovered and reported many interesting cases of synæsthesia particularly number forms and colored vowels. His work attracted wide attention to the subject and stimulated a large number of contributions to the literature. Case after case could be cited, but these would add little, if anything, to the solution of the problem proposed for investigation. Anyone interested in reading descriptions of such cases can readily locate them by reference to the extensive bibliographies available. That of Mahling (**16**) contains 550 titles and Argelander (**2**) gives 466 titles.

We shall now review briefly the various theories that have been advanced as an explanation of this 'curious' phenomenon. Before reviewing the theories, however, let us consider the following facts regarding chromæsthesia which the theories seek to explain.

In the majority of cases, chromæsthesia seems to be related to a high degree of imagination. Rarely is there a defect of either vision or hearing, and in a number of cases, these senses have been especially keen. For instance, Nussbaumer was able to distinguish eleven partial tones without the use of resonators and without practice. With most persons the color-sound associations date back to very early years; many individuals are unable to remember when they did not have such tendencies. Sound, especially speech and musical tones, are most frequently linked with color, but in some cases, letters and whole words are colored. The two sensations are usually inextricably linked and appear simultaneously. There is no measurable difference in the reaction time for responding to the linked sensations.

The tendency toward chromæsthesia in some cases seems to be hereditary. In almost all of the cases reported, other individuals of the same family have had similar tendencies.

Lundborg (14) has recently investigated colored hearing in three generations of a Swedish family and concluded that it is probably a dominant Mendelian factor but not sex linked. In spite of the evidence in favor of an hereditary concept of chromæsthesia, such a theory must be broad enough to explain wide idiosyncracies, since, as Galton (7) states, "No two people agree, or hardly ever so, as to the color they associate with the same sound." Certainly these facts indicate the influence of experiential factors which must be considered in any adequate theory.

Other facts to be explained are these: fatigue seems to facilitate the color impressions; certain drugs may artificially produce these phenomenon; there is a remarkable constancy of specific color-sound associations over long periods of time; the intensity and clearness of the colors may be different for the right and left eye; and practically all cases agree that it is the sound itself, and not the meaning, which calls up the sensation or image of colors.

Although many possible explanations of synæsthesia have been advanced by various writers, they seem to fall into two main groups: physiological and psychological. In general, it may be said that the physiological theories also assume an hereditary view of synæsthesia, while the psychological ones seek to explain the facts on an environmental or experiential basis, but this is not always the case. Certain physiological theories also belong to the environmental group. The reverse, however, does not hold. All psychological theories are stated in terms of past experience.

It must be admitted that none of the proposed theories, either physiological or psychological, satisfactorily explains all of the facts mentioned above. In the first place, any theory of chromæsthesia necessitates the acceptance of one of the various theories of color vision, none of which have received universal acceptance. Assuming, however, that a satisfactory explanation of color vision were available, there is still need for a suitable explanation of why two or more types of sensations are aroused by stimuli which are ordinarily perceived in only one sense field. Let us consider first the physiological theories that have been advanced.

The earlier writers considered cases of chromæsthesia as pathological in nature. Carnaz in 1848 (**3**) explained them as due to a hyper-æsthesia of the sense of color, probably due to some optical lesion. Lussana (**15**), in 1873, suggested that the brain centers for perceiving color and sound might be contiguous, and thus have reciprocal influence on each other. Nuel (**17**), agreeing with this view, suggested that these secondary sensations might be due to nervous irradiation of sensorial currents. Pouchet and Tourneux (**20**), with others, have held that chromæsthesia results from abnormal crossing or even anastimosis of afferent nerve fibers. Fêrè (**6**) considered the phenomenon to be due to a particular tonus of the nervous system, produced either by drugs or an unusual experience. The action of hasheesh and mescal, already referred to, offer a certain amount of evidence in favor of this theory.

A novel physiological theory of chromæsthesia as well as of synæsthesia in general was suggested by Titchener (**23**), who explained the tendency toward double sensations as resulting from an unusual elasticity of the walls of the cerebral arteries. A rush of blood to the auditory center might also affect the visual center, thus resulting in colored hearing. While this theory is interesting and fits in nicely with certain known facts concerning chromæsthesia, such as the loss or reduction of the tendency with increasing age and its occurrence in moments of emotional stress, it nevertheless does not offer a very satisfactory explanation of the phenomenon. Perhaps, its most obvious limitation is that the relatively slow change in blood pressure would not account for the practically simultaneous appearance of the double sensations.

While considering the nature of synæsthesia in general and chromæsthesia especially, it occurred to the present writer that these phenomena might be nothing more than examples of the well known conditioned reflex. In surveying the literature, it was discovered that Wheeler (**27**), after an extensive investigation of the synæsthesias of a blind subject, had ventured to define synæsthesia as 'an immediate and permanent conditioned reflex.' He admitted that such a

definition was purely hypothetical and offered it merely as being as good or better than any previously proposed theory. He failed, however, to suggest ways and means of producing an 'immediate and permanent' conditioned reflex, or to specify why such might occur in synæsthetic subjects.

Many investigators have felt that a physiological explanation of synæsthesia is impossible and that a psychological one is needed. While most of the theories advanced by this group are based on the law of association of ideas, they differ in certain respects, and will be reviewed briefly.

Kaiser (**11**) suggested as an explanation of a case of 'colored seeing' that the subject himself puts the color into an intimate association with the words in order to engrave them better upon his memory, the association extending back to early childhood, and while it is at first voluntary, it later becomes spontaneous. Stevens (**22**), Schenkl (**21**), and others have offered similar explanations, holding that the subject either intentionally or accidentally associates the two types of sensations and thereafter, the one arouses the other. Such theories are not by any means extinct. In his recent text, Woodworth (**28**) offers only the following brief explanation of the phenomenon: "The synæsthesia probably resulted in childhood from playfully linking the sensations together." Evidently some persons have highly developed powers of introspection—even in childhood!

To the present writer, the psychological theories which have been proposed seem superficial and inadequate. According to Titchener (**23**), "Attempts have been made by trained and interested observers to trace back their synæsthetic experiences to associations formed in childhood; but in spite of all their efforts, they have ended in failure." The suggestion that the phenomenon is an example of the conditioned reflex represents an attempt to offer a combined physiological and psychological theory, that is, to explain the known facts of psychological association on a neurological basis. The fact that the order of synæsthesias are never reversed forcibly suggests the analogous unidirectional action of the conditioned reflex.

As mentioned above, almost all of the extensive literature bearing on the subject of chromæsthesia is either descriptive or anecdotal. There has been much writing and theorizing concerning the subject, but very few experimental studies have been carried out to clear up the points in question. The small amount of experimental work that has been done may be divided into three general groups: (1) repeated observations (with the aid of controlled auditory stimuli) on individual cases of chromæsthesia; (2) group experiments with unselected populations using controlled auditory stimuli with the object of establishing qualitative laws between sensations of tone and color; and (3) so-called statistical studies or surveys which have sought to determine the prevalency of chromæsthetic tendencies. It should be pointed out, however, that with the exception of the first type of study mentioned, the investigators were studying associations of sensations and not the spontaneous double sensations which are found in marked cases of chromæsthesia. Surprising as it may seem, apparently none of the many investigators who have been interested enough in chromæsthesia to contribute to the literature of the subject, have made any attempt to establish this tendency toward colored hearing in normally non-synæsthetic subjects. Off-hand, it would seem that at least some of the many theorists would have subjected their theories to such a crucial experimental test. This being the case, the writer felt that it would be desirable to conduct an experimental attempt to establish artificial chromæsthesia under controlled conditions.

II. Technique of the Experiment

Inasmuch as many students of chromæsthesia have held that the sensations are the result of associated past sensory experiences, and since the conditioned reflex theory presupposes a similar association of stimuli, except that they must be experienced under certain rigid conditions, it was decided to make the experimental situation such that it would determine the possibility of chromæsthesia resulting either from simple association or from conditioning, thus testing both of

the theories at the same time. At this point, the writer wishes to emphasize the fact that he had no personal interest in either theory. He merely wanted to put the hypothesis to an experimental test. The success of previous investigators in establishing conditioned reflexes in human subjects led him to believe that it might be possible to condition an artificial chromæsthesia.

As is well known, the necessary conditions for the formation of a conditioned reflex are two (**19**): (1) there must be an existent reflex, which may be either a conditioned or an unconditioned one; and (2) the conditioned and unconditioned stimuli must occur either simultaneously or the conditioned must precede the unconditioned one. It would seem that these requirements could be met easily, since (1) there is an existent reflex or response, that of color stimuli eliciting the sensation of color; and (2) it is a fairly simple matter to arrange for the simultaneous presentation of the conditioned and unconditioned stimuli (tone and color pairs). The possibility of meeting the latter is certain since it obviously entails nothing more than overcoming technical difficulties in the devising and building of the necessary apparatus. Satisfying the first, however, that of utilizing an existent reflex or response is a somewhat more questionable possibility.

Does the act of perceiving a color involve either a reflex or a response? We are face to face with the unanswered question of the nature of sensation and with the problem of where sensations are perceived. To say with the structuralists that sensations occur in consciousness does not suggest any possible clue as to where they take place or as to what physiological or neurological activity is involved. Physiology and neurology have likewise failed to answer the question in an adequate manner. To be sure, we know that certain end organs, nerves, and brain areas are necessary for the conscious perception of sensation, but the exact mechanism involved is as yet undetermined. It does appear, however, that some type of a response takes place, and upon this assumption, the validity of the present experiment must rest. Recently Allen (1) has maintained that afferent sensory impulses not

only arouse characteristic sensations in the cortical centers, but furthermore arouse certain efferent currents or reflexes which either heighten or depress the sensitivity of their receptors. In view of the unknown nature of the assumed response, however, the present problem is hereafter referred to as an attempt to establish a *conditioned response* rather than a *conditioned reflex*.

This brings us to one of the most interesting theoretical aspects of the investigation. If it is shown that it is possible to establish a conditioned response upon the act of perceiving a sensation as the unconditioned response, a valuable contribution will have been made to our understanding of the true nature of sensation. On the other hand, if under the proper conditions, such conditioning fails to occur, this will constitute one line of evidence that sensations are neither reflexes or responses, or at least not the ordinary type of reflex or response which has been shown amenable to conditioning.

As a part of the experiment, it seemed advisable to make use of a number of different stimuli, both conditioned and unconditioned. For this purpose, it was decided to use the eight notes constituting the C major diatonic scale each paired with a different color. Essentially then, the experiment represented an attempt to establish a number of conditioned responses at the same time. In order to make sure that this complication could not be the cause for a possible failure of the experiment, it was also decided to devote a part of the experiment to single tone-color pairs and thus determine the possibility of establishing a simple crossing over of the two sensations.

In still another respect, this experiment differs from other conditioning experiments, in that neither the conditioned nor the unconditioned stimuli used appeal to any strong physiological drive such as sex or hunger. It is conceivable that it might fail because of this fact were it not that Cason's (**4**) experiments with the pupillary and eyelid reflexes were successful in spite of a seeming lack of motivation.

In order to take advantage of any possible existent relationship between color and tone (generally known as the

Newtonian parallel) it was decided to pair each note of the octave with the color occupying the same relative position in the spectral series as the note occupies in the musical scale. Practical exigencies made it desirable to use only six different colors in addition to white and these were paired with the notes in the following manner:

Note	Color
C	White
D	Red
E	Orange
F	Yellow
G	Green
A	Blue
B	Violet
C^1	White

Theoretically, for fulfilling the conditions necessary for the building up of a conditioned reflex, any color might have been paired with any tone. However, since it was possible to meet the required conditions and also retain the order suggested by the Newtonian parallel, it was considered wise to do so.

The pairing of white with both the upper and lower 'C' calls for a word of explanation. No one has reported any relationship between the colors accompanying octave pairs in actual cases of chromæsthesia. Since the frequencies of such pairs are always in the ratio 1 : 2, it seemed logical that if conditioning did occur, and if any color at all was seen upon the presentation of a tone one octave higher, it should be the same color as elicited by the same note an octave lower. In case the hypothesis did not prove correct, this procedure could not invalidate the entire experiment since six other different color-tone pairs were used.

Another interesting theoretical aspect of the experiment was the possibility of its revealing new facts regarding color vision. Let us assume for the moment that conditioned responses were established for each of the above tone-color pairs. According to the pairing, the subject would see yellow upon hearing the tone 'F' and blue upon hearing the tone 'A.' What color would be seen if both 'F' and 'A' were sounded simultaneously? Would it be green, grey, or a

motley mixture of the two, or even something else? The answer to this and similar questions which might be asked would be exceedingly useful to students of visual and optical phenomena.

Apparatus

The essential demands of an adequate experimental set-up for this problem were as follows:

1. It must include apparatus capable of producing sustained tones corresponding to the eight notes of the C-major diatonic scale.

2. It must provide for presenting any one of seven colors of light *all in the same place*, in order to prevent any spatial orientation of the colors.

3. It must be so designed and constructed that certain tone color pairs will always be presented simultaneously.

4. It must be automatic and operate at a constant speed in order to assure consistency of the stimulus series.

5. It must be practically devoid of extraneous noises.

These requirements were all fulfilled in the set-up which was constructed for the experiment. Specifically, they were met as follows:

1. After much casting about for a suitable means of presenting the sustained tonal stimuli, it was at last decided to use a 'grind organ' such as was popular in this country several years ago. A thorough search failed to discover a suitable instrument of this type and other possibilities were considered. At length, the writer decided to construct a 'grind organ' by making use of a piano keyboard accordion, an electric blower unit,[2] and a roller device for making it 'self-playing.'

2. At first it was planned to use eight colored electric light bulbs, arranged in a row, as the unconditioned stimuli. It was evident, however, that such an arrangement would result in a spatial orientation of the colors. Since this would have added an unwanted factor to the unconditioned stimuli, a possible means of presenting all of the colors in the same place was sought. This was accomplished by the use of seven projection lanterns, each bearing a colored gelatin filter and all focused on the same area of a white screen. In order that each square of color would appear to be exactly the same size and shape and in exactly the same location as all others, a white square was laid off on the screen and bounded by dull black cardboard.

3. The simultaneous presentation of the tone-color pairs was accomplished by means of electrical contacts soldered to each of the eight keys of the accordion. Each time one of these keys was depressed to produce a tone, the contacts closed; this in turn operated a relay which lighted the projection lantern carrying the color of filter with which the tone was paired.

4. The apparatus was made to function automatically by means of a revolving cylinder carrying studs which depressed each of the keys in a prearranged order. Constant speed of operation was assured by the use of an electric phonograph motor geared to the roller.

5. Preliminary experimentation showed that the only noisy parts of the set-up were the blower unit and the relays. In order to eliminate these distracting noises, the blower was placed in a room on the floor above and the compressed air led to the ac-

[2] A used vacuum cleaner motor, fan, and housing served admirably for this purpose.

cordion by means of a one-inch pipe, and the relays were placed in a room adjoining that which contained the tone unit and the projectors.

With the exception of the blower unit and the relays, the entire set-up was arranged in a room 11 feet wide and 15 feet long. The screen was placed directly in front of the tone unit, hence the tones seemed to emanate from the colors projected on the screen, the illusion being similar to that of the talking movies. The seven projection lanterns were mounted on a table which was elevated to a sufficient height to throw the colors over the heads of the subjects who were seated in the room.

In order to record the number of repetitions of the paired stimuli, a small Veeder counter was connected to the playing cylinder. The same movement of the rod which actuated the counter was used to operate an automatic stopping device which could be set to shut off the power supply and thus stop all of the apparatus at the end of from 1 to 30 revolutions of the playing cylinder. This device permitted the experimenter to serve also as a subject, since it relieved him of the necessity of counting the number of stimuli presented and of stopping the apparatus at the proper time.

In order to make possible the continued repetition of a single tone-color pair, a 'magnetic finger' was constructed, which could be used for depressing any desired key. This finger was controlled by means of a five segment commutator mounted on the shaft of a synchronous electric motor. A second automatic switch was used to stop this motor at the end of ten minutes or after 50 repetitions of the stimulus pair. For the individual tone-color pairs produced by the magnetic finger, the duration of each stimulus was approximately 9 seconds with a three second interval between successive stimuli.

The rate and also the duration of the presentation of the entire scale of tone-color pairs was determined by the speed of the electric motor driving the playing cylinder. For the purpose of the experiment, it was adjusted so that this cylinder made one complete revolution in 45 seconds. Since 16 pairs of stimuli (one ascending and one descending scale) were produced by each revolution and since there was an interval of approximately 2 seconds between each scale and one-half second between each stimulus pair, the actual duration of each tone-color pair was approximately 2 seconds.

It has been determined that conditioned responses are established when the conditioned and the unconditioned stimulus occur simultaneously, or the conditioned stimulus precedes the unconditioned, but not vice-versa, hence the apparatus was adjusted so that the tones appeared either simultaneously with or else just before the colors. Since the keys of the organ had to open slightly before the electrical contact was closed, the tones were produced as soon as, or slightly before the colors were thrown on the screen.

The apparatus, from the mechanical and electrical point of view, functioned perfectly throughout the course of the experiment.. It is felt that it not only provided adequate stimuli, but also met all other requirements for carrying out the experiment in a satisfactory manner. Psychologically considered, the tonal stimuli were of course not pure tones, since each note of an accordion is produced by two vibrating reeds each accompanied by its characteristic overtones. However, the presence of these secondary cues only served to make tones more easily distinguishable and therefore all the more desirable. Likewise the color filters were not chosen with regard to their pureness of color but rather for their readily distinguishable nature.

III. Experimental Procedure

The experiment itself was in part a group experiment and in part an individual experiment. The group experiment constituted an attempt to establish chromaesthesia for an entire octave of the musical scale; the individual part consisted in the attempt to condition one tone to one color only.

In the group experiment, the subjects were seated in front of the screen and all lights extinguished. In order to assure a similar mental set on the part of all subjects, the purpose and technique of the experiment were explained before the experiment began. The only instructions given to the subjects were: "Look at the colors and listen to the tones," but it was specifically stated that the subjects might assume any attitude desired, or use any method they wished in reacting to the stimuli. In the group experiment, the stimuli were presented in series of 40 complete scales, 20 ascending and 20 descending ones. Two such series or 80 repetitions of each of the 8 tone-color pairs were presented at each sitting. The group met four times per week: Monday, Tuesday, Wednesday and Thursday at 5 P.M., so that a total of 320 repetitions were presented weekly for 7 weeks.

For the individual experiment, the subjects were each assigned a single tone-color pair. These subjects were instructed in the operation of the apparatus, and at times convenient to themselves, entered the dark room individually for their part of the experiment. They, too, sat directly in front of the screen, and followed the same instructions as the subjects in the group experiment: "Look at the color and listen to the tone." As stated before, the individual tone-color pairs were presented at the rate of 5 per minute, or a total of 50 during the ten minute period constituting each series. Two or more series of 50 pairs of stimuli were often presented at the same sitting according to the amount of time the subject could devote to the experiment on that particular occasion.

The total number of presentations of the paired stimuli which would be required to establish the conditioned response, if at all, was of course unknown. Pavlov (**19**) and his co-workers have succeeded in conditioning dogs with as few as five repetitions of the stimuli, while the greatest number of repetitions required in any previously reported experiment was 400 in the case of the pupillary and eye-lid conditioning experiments of Cason. For the present experiment, it was felt that at least 2000 repetitions of the paired stimuli should be presented if a conditioned response was not established with fewer. It was, of course, almost a physical impossibility to present all of the 2000 stimuli at one sitting. Since it has been shown that interruption even for periods of months does not destroy a conditioned reflex, it did not seem that the presentation of the stimuli on successive days would seriously invalidate the experiment. In order to make sure that the spaced stimulations were not responsible for a failure of the response to become conditioned, the experimenter and another subject taking part in the group experiment subjected themselves to 1000 successive presentations of the eight pairs of stimuli at one sitting which lasted six and one-half hours. A similar test of the relative effect of spaced and concentrated stimulation was made for a single tone-color pair. Here again the experimenter and one other subject sat through 1000 repetitions of the stimuli. In this case the sitting lasted about three and one-half hours.

Subjects

A total of 18 subjects took part in the experiment. All were university students, mostly graduate students, and all had had one or more courses in psychology. Of the

18, thirteen were members of the group experiment utilizing the entire scale. The other five took part in the individual experiment with single tone-color pairs, and each was presented with at least 1000 repetitions of the paired stimuli. Since the group experiment required an exhorbitant amount of time from each subject and involved the inconvenience of meeting over a period of seven weeks, it was not possible for all of the 13 members of the group to continue through the course of the experiment. However, five of them were presented with at least 2000 repetitions of the stimuli, and one, the experimenter, experienced over 3000 stimulations. The remaining eight subjects were presented with from 280 to 1000 stimulations.

A series of preliminary tests and a questionnaire revealed the facts that: all of the subjects had normal color vision; all had an intrinsic interest in the problem; all were sceptical of the possibility of positive results; none of the subjects had any natural tendency toward chromaesthesia but one subject was synæsthetic to the extent of having number and date forms; and that the group represented an approximate chance selection with regard to artistic and musical ability. The results of a color and tone association test confirmed the above statement that the group was primarily non-synæsthetic.

Method of Ascertaining Results

The experiment as outlined above provided for two types of results: (1) an answer to the question of whether or not it is possible to produce chromæsthesia by the technique of the conditioned response (or whether chromæsthesia may result from the simple association of color and tone); and (2) the indication of any association learning which might have occurred. The major purpose of the experiment was, of course, to determine the possibility of conditioning chromæsthesia. While it was expected that a certain amount of associational learning would occur, this aspect of the experiment was incidental.

Since the experimental situation was such that both conditioning and learning might occur, it is necessary to set up certain criteria by which to interpret the results obtained. Although Pavlov (19) and others regard learning and conditioning as one and the same thing, the writer believes himself justified in making a distinction between the two processes in this case. It was to be expected that the simultaneous presentation of the stimulus pairs would result in a type of abstract learning of the paired tones and colors. Since only eight pairs of stimuli were used, it would have been but a simple matter for anyone with a good sense of pitch to memorize the colors corresponding to each of the tones. Since the tones were easily distinguishable by most of the

subjects, it would have been surprising if they did not acquire the ability to name the color paired with any given tone. Obviously, this could not be taken to indicate the formation of a conditioned response, but merely the learning of paired associates.

By definition, the formation of a conditioned response requires that the substitute or conditioned stimulus become an adequate stimulus for eliciting the response formerly called forth by the unconditioned stimulus. In the present experiment, this response consisted in the seeing of colors. If the tones became capable of arousing spontaneous sensations or images of color in the subjects, then and only then, could it be said that a true conditioned response had been established.

Because of the non-overt nature of sensation, it was practically necessary to utilize the verbal reports of the subjects in determining the results of the experiment. In spite of the unreliability of introspective reports in general, it is believed that those obtained in this study were both valid and reliable. The subjects were all fairly mature and cognizant of the need of accurate reports. As was evidenced by their intensive and loyal coöperation, they were all interested in the outcome of the experiment. The type of introspection called for, that is, the mere report of the presence or absence of color sensation was extremely simple, and lastly, there was no reason to report falsely.

The method used for obtaining these verbal reports also permitted the measurement of the amount of associational learning that had occurred. Briefly, it was as follows: at the end of each Thursday evening session, each subject was provided with a copy of a mimeographed sheet containing 40 numbered blank spaces. The following instructions appeared at the top of the sheet and were read aloud by the experimenter:

Directions.—You will be presented with a series of forty tones made up of notes picked at random from the scale. After each note is sounded, you are to write in the color which goes with that tone. Use the first letter rather than the entire word for the names of the colors. If you have a real image or sensation of the color, put your response in the column marked 'color'; if you do not seem to see the color but think that you know what color goes with the tone sounded, put your response in the column

marked 'know'; and finally if you neither see the color nor know what color goes with the tone, guess a color and write the name of the response in the column marked 'guess.' Do not omit any response. If you do not know, guess!

Following the reading of the instructions, a series of forty tones were sounded in random order. Each of the eight notes were thus presented five times. The subjects were asked to guess if they did not know the correct response on the assumption that a certain amount of learning might have occurred without the subject being aware of it. Since all subjects guessed when they did not know the correct answer, goodness of performance on each test could be expressed in terms of the number or per cent of correct responses and thus permit direct comparisons. This method also made possible the drawing of both individual and group learning curves.

Results

The results of the experiment with regard to the conditioning of chromæsthesia were entirely negative. Not one of the eighteen subjects acquired any tendency to see colors upon the presentation of the auditory stimuli. There was not a single response in the column marked 'color' on any of the test blanks. Indeed the subjects' interest in the outcome of the experiment was so great that had anyone of the subjects experienced a real color sensation on hearing the tones, it is likely that he would have announced the fact at once rather than waiting to write down the response in the proper column. Conditioning, in the sense that has been described above, simply did not occur.

With regard to the learning aspect of the experiment, an analysis of the subjects' performance on the weekly 40 item test indicated that something analogous to learning had taken place. The number of subjects taking each test,[3] the number of previous presentations of the stimuli, and the mean number of correct responses to the test are shown in Table 1. As will be readily seen, there is a tendency for the correct number of responses to increase with the number of repetitions

[3] It was, of course, impossible for the experimenter as a group subject to both administer and take these tests.

TABLE 1

RESULTS OF THE CHROMÆSTHESIA EXPERIMENT

Test No.	1	2	3	4	5	6
No. of stimuli	360	680	1000	1320	1640	2000
No. of subjects	8	9	7	4	4	4
Mean score	22.5	23.1	28.0	26.0	31.0	31.7

of the stimuli. In other words, a certain amount of learning did occur. Since there were 40 responses on each test and only seven different colors were used, mean scores of 5.7 would have resulted from sheer guessing. Since the means of all the tests, even of the first one, are considerable over this value, it indicates that a definite amount of learning did occur. When plotted, the graphs for both the group and for the individual members, resembled the typical learning curve.

Certain factors, however, make it doubtful that the increase in performance can be attributed solely to learning. For instance, the scores on test 1 ranged from 12 to 38 correct responses out of a possible 40. Even on the last test, after 2000 repetitions of the stimuli, the four scores were 22, 32, 33, and 40. Since there were only eight pairs of stimuli involved, it is obvious that something other than a lack of learning ability was responsible for the failure of all subjects to make perfect scores after the first few repetitions of the stimuli. The most likely factor was the inability to identify the eight tones. Such being the case, it is conceivable that the increase in ability to name the corresponding colors might be attributed to an improvement in the sense of absolute pitch. However, an absolute sense of pitch was not necessary to recognize the different tones, since as has been mentioned, certain secondary cues were present in the tonal stimuli.[4]

That the ability to name the colors corresponding to the various tones was not simply the result of learning is supported by the fact that the subject, who with the experimenter, sat through 1000 repetitions of the 8 tone-color pairs was able to respond correctly to but 17 tones out of a series of

[4] This suggests the desirability of conducting an experiment utilizing pure tones and paired stimuli to determine the possibility of improving the sense of absolute pitch.

40 presented at the end of the session. In this connection it is interesting to note that the experimenter, after having been subjected to over 3000 repetitions of the stimuli, 1000 of which were consecutive, and after having planned and executed the experiment, was able to respond with the correct color but 25 times out of 40. What could be better proof that he had not become chromæsthetic?

Supplementary Experiments

Although there was not the slightest evidence that any conditioning had occurred even after the thousands of repetitions of the stimuli, certain supplementary experiments seemed desirable before concluding that such conditioning was impossible. It was conceivable that a weak connection had been established which was not sufficiently strong to cause the appearance of introspectively observable colors upon the presentation of the sound stimuli. Two possible methods of facilitating or augmenting any such latent tendency seemed worthy of trial.

As was mentioned earlier in this paper, Fêrè (**6**) held that colored hearing was due to the particular tonus of the nervous system, produced either by drugs, fatigue, an unusual experience such as a bad fright, or other emotional disturbance. Although this opinion is not widely held, it was thought worthwhile to carry out the following simple experiment.

As on several previous occasions, the experimental room was darkened and the subjects told that two complete ascending and descending scales would be played without the colors being presented. They were instructed to gaze intently in the direction of the screen while noticing carefully for the appearance of any possible color sensations or images that might occur. Between the first and second series of tones, a blank pistol cartridge was fired. Since the shot was entirely unexpected by the subjects, this loud report produced a most admirable emotional disturbance, if the reactions and reports of the subjects were a valid indication. In spite of the disturbance, none of the subjects observed any color sensations upon hearing the series of tones presented immediately after the shot.

Likewise fatigue did not facilitate any tendency toward chromæsthesia. On several occasions, the writer listened to the tones late at night after an unusually tiresome day, and yet there was not the slightest appearance of color. As can well be imagined the writer and the subject who sat through the six and one-half hour session of 1000 consecutive stimulations were in a state of fatigue, and yet in spite of both the recency of the stimulation and the fatigued condition, color sensations did not result from hearing the auditory stimuli.

The use of certain drugs as facilitating agents seemed somewhat more promising. It has already been noted that Gautier (**8**), in 1843, reported that he had been able to produce temporary chromæsthesia by the use of hasheesh. Havelock Ellis (**5**) observed that one of the effects of the drug, mescal, was to produce a tendency toward synæsthesias of various kinds. In view of these and other reports, it was decided to study the influence of mescal in facilitating any possible 'latent chromæsthesia' that might have resulted from the extensive conditoning process. Since mescal [5] produces both colored imagery and, at the same time, effects a temporary tendency toward chromæsthesia, it was thought that it might enhance any basic tendency that had resulted from the conditioning process so that the connection would be observable in the imagery of the subjects. If any colors were seen when under the influence of mescal, and if any connection had been established between the tones and the colors, it was expected that the paired colors would appear on the sounding of the various tones. It seemed also possible that the colors might be aroused by the tones before they began to appear spontaneously. And finally, it seemed likely that in view of the enormous number of stimulations by colored squares, there would be a predominance of squares among the colored imagery seen by the subjects under the influence of the drug.

Because of the many unpleasant effects of the drug, the experimenter planned to take the mescal himself rather than to ask any of his subjects to do so. However, upon informing

[5] For an excellent account of the properties and psychological effects of this interesting drug see: Kluever, Heinrich, *Mescal: The Divine Plant and Its Psychological Effects*, pp. 111. Psychic Miniatures, No. 22. Kegan Paul, London 1928.

the four subjects who had continued throughout the experiment of his plans, he was surprised to find that they also wished to take part in this phase of the experiment. Accordingly, it was planned that the entire group of five should take mescal. Another engagement prevented one of the subjects being present on the day chosen for the experiment, so actually only three of the group subjects in addition to the experimenter took part in this procedure. It seemed desirable also that at least one of the five subjects who had been presented with single tone-color pairs should also take part in the mescal experiment. Luckily, the subject who with the experimenter had been stimulated with 1000 consecutive repetitions of the single tone-color pair (A-blue) volunteered his services in this connection. The actual taking of the mescal was carried out just one week after the last meeting of the experimental group and on the day following the just mentioned intensive study of the single tone color pair. It seemed likely, therefore that the effect of sitting through three and one-half hours of stimulation by the blue squares would surely be noticeable in the resulting imagery of the individual subject and the experimenter.

In all then, five subjects including the experimenter actually took part in the mescal experiment. Although standard doses of the drug had not been determined, the reports of previous investigators indicated that six buttone, or approximately 15 grams, would probably be sufficient to produce the desired imagery and yet not enough to be dangerous to life or health. The buttons were mashed with mortar and pestle to make ingestion reasonably easy, and were then swallowed with the aid of water. Three doses of 5 grams each were taken with the aid of about 8 ounces of water. The first dose was administered at 2:10 P.M., at which time each subject was provided with a notebook and instructed to keep a diary of his experiences. At 2:45 the subjects entered the dark room and a series of tone scales were presented. The second dose was administered at 3:10 followed by a second presentation of the tonal stimuli at 3:45 while the third and last dose was given at 4:10. Additional

tests of the adequacy of the tones for eliciting color imagery were made at 5:15 and again at 7:45 P.M.

Inasmuch as the detailed description of the reactions and imagery of the five subjects are of no particular value in connection with the present investigation, they will not be reported here.[6] In the present connection, suffice it to say that all but one of the subjects were rewarded for their pains by gorgeous arrays of colored visions which compared favorably with those described by previous investigators of the so-called divine plant. For all but one, the shows were subjectively expensive, since all of the subjects except the experimenter suffered severe attacks of nausea. Unfortunately, the one subject who did not experience unusual imagery of any kind seemed to suffer more than anyone else. Even a fourth dose of 5 grams taken at 5:10 failed to produce any unusual imagery for this subject.

Here again, there was no evidence that any connection had been established between the colors and tones used in the experiment. In spite of the fact that a constant train of colors was passing in the imagery of the subjects, the tonal stimuli not only failed to arouse the colors with which they had been paired, but also failed to alter materially the imagery being experienced at the time. Likewise, the tones failed to arouse colored imagery before it began to appear spontaneously. Either the conditioning process had failed to build up a basic framework of chromæsthesia or else mescal did not facilitate the connections sufficiently to make it introspectively observable. The latter alternative seems unlikely, since the mescal did produce certain synæsthetic tendencies in all of the subjects reacting positively to the drug. The most marked combinations of sense fields seemed to be between vision and touch and between vision and kinæsthesia, although there was some tendency for all types of sensations to be colored. Two of the subjects even remarked about their colored pains.

The failure of the conditioning process to influence the mescal imagery was further indicated by an almost complete

[6] Kelly, E. L. *Individual Differences in the Effects of Mescal.* Jour. General Psych. 1933, 9, 462–472.

absence of colored squares in the imagery experienced by the subjects. Even the experimenter, who enjoyed an unusually beautiful array of colored images, observed no indication that the 1000 stimulations with the blue square on the previous day influenced his imagery in any way. There was simply no observable evidence of any kind that the subjects had engaged in the experiment for the past seven weeks. Thus we find that the use of mescal served only to emphasize the negative nature of the experimental findings.

IV. Interpretation of Results

We have seen that in spite of the extensive conditioning procedure and the seeming adequacy of the methods used to check the results of the experiment, there was not the slightest evidence to indicate that a conditioned response had been established. Artificial chromæsthesia was not produced in any of the subjects, and the mescal experiment failed to reveal any connections between tonal and visual stimuli used. As was stated in the beginning of this report, positive and negative results were equally desirable. Without question, the results turned out negative. They are so distinctly so that the writer has no hesitancy in concluding that it is impossible to produce chromæsthesis in normally non-synæsthetic adult subjects by the technique of the conditioned response.

A number of reasons might be suggested as to why a conditioned response of chromæsthesia was not established. First, it might be argued that the number of stimulations was not sufficiently large. However, in view of the fact that conditioned responses have been established by as few as five presentations of the stimuli and that in other cases it has not been necessary to use more than 400 repetitions, the thousands of stimuli used in this experiment seem unusually adequate. It is very unlikely that conditioning would have resulted had a larger number of stimulations been presented.

It might also be suggested that the spaced conditioning periods prevented the formation of the conditioned response. It will be remembered, however, that two subjects were given 1000 consecutive stimulations with no different results.

Again, it might be thought that conditioning failed to occur because of the inhibitory effect of the several pairs of stimuli used. This can hardly have been the case, since five subjects were presented with only a single tone-color pair, and there was no more evidence of conditioning than when the eight pairs of stimuli were used.

While other minor objections might be raised in regard to the experimental technique employed, the procedure appears to have been so nearly adequate that the reason for the failure to establish a conditioned response must be sought elsewhere. It will be recalled that the two requisites for the formation of a conditioned reflex (response) are: (1) There must be an existent reflex (response) which may be either an unconditioned or a conditioned one; and (2) the conditioned and unconditioned stimuli must occur simultaneously, or the conditioned must precede the unconditioned stimulus. Unquestionably the second condition was satisfied in the present experiment. On the other hand, it will be remembered that it was only assumed that we were dealing with an existent response, inasmuch as the seeing of color seems to be a response to the light waves which are perceived as color. The questionable nature of sensation has already been mentioned, and the Weiss-Warren controversy (**24**, **26**) of a few years ago served only to emphasize the fact that we do not know whether sensation involves a somatic response as maintained by Weiss, or as Warren holds, consists only of a neural disturbance ending in the central nervous system.

A possible explanation, therefore, for the failure to establish a conditioned response of chromæsthesia is that there was no existent response to become conditioned. In other words, the negative results of the experiment might indicate that the perceiving of color is not a response. This interpretation fits in nicely with Warren's view of the cortical nature of sensation, and it seems to the writer to be the most logical explanation for the failure to obtain positive results in the present experiment. At least, we can conclude that sensation is not a response of the ordinary type which has been shown to be capable of being conditioned.

The bearing of this conclusion on the conditioned reflex theory of chromæsthesia is evident. If 3000 successive presentations of the tone and color stimuli failed to produce chromæsthesia in normally non-synæsthetic subjects, it can hardly be assumed that the chance simultaneous occurrences of visual and auditory stimuli in actual life could result in the marked cases of colored hearing which are occasionally encountered. Likewise the findings indicate the gross inadequacy of the psychical association theory of chromæsthesia.

Certain qualifications of these conclusions must be made. The present experiment showed that it is impossible to condition chromæsthesia in a group of adult subjects. It is conceivable, though hardly likely, that it might be possible to produce chromæsthesia in more youthful subjects by the conditioned response technique. It may be that such stimulations as occur during early childhood or under certain emotional strains may result in the cases of chromæsthesia which occur. The results of the present study, therefore, do not permit an all-inclusive denial of the conditioned reflex theory of chromæsthesia. It seems extremely desirable that a similar experiment be conducted with very young subjects and with subjects under severe emotional strain, in order to test the validity of the technique under these conditions.

Inasmuch as linkages between colors and tones of differing timbre are reported more often than linkages between colors and tones of specific frequencies it is conceivable that conditioning might have resulted from the use of tones widely different in quality, such as produced by different musical instruments. It does not seem likely that such positive results would have been achieved by these means, since there was no indication of conditioning even in the experiment employing single tone-color pairs.

For still another reason the writer holds that the conditioned reflex and psychical association theories are not adequate explanations of chromæsthesia. If, as they assume, the mere simultaneous occurrence of certain stimuli result in a tendency toward later double perception, why is it that marked cases of chromæsthesia are so rare? Certainly a large

proportion of growing children are exposed to many simultaneously presented stimuli, yet only a very small number of cases of true chromæsthesia result. It seems, therefore, that there must exist, in a few persons at least, a predisposing physical or physiological basis which permits connections to be readily formed. Such a physical or physiological disposition might be either inborn or the result of somatic changes occurting during the life of the individual.

While the writer does not wish to propose a new theory of chromæsthesia, he does feel that some such hypothesis as the above must be incorporated in anything like an adequate explanation of the phenomenon. By assuming a physiological disposition plus a conditioning process, most of the known facts of chromæsthesia can be explained readily. The hereditary tendency of the trait as well as the wide idiosyncrasies which have been reported are thus reconciled in one explanation, a thing which all other theories have failed to do. Likewise the rarity of cases of chromæsthesia can be explained on the assumption that suitable predispositions are rarely existent upon which the conditioned response may become established.

The negative results of the present experiment are also quite explainable by such a theory. If the physiological predispositions necessary for the formation of the conditioned responses are rare, as is suggested above, it is not likely that one would have appeared in the relatively small group of subjects used in the experiment.

It has been argued by some that a purely physiological theory is sufficient without including the assumption that conditioning occurs. A crucial test of whether conditioning does supplement a physiological factor in the production of chromæsthesia could be made by attempting to destroy and change the color-tone associations in *bona fide* cases of colored hearing. If it could be shown that it is possible to experimentally destroy the linkages reported, it would offer almost incontrovertible evidence that conditioning had played a rôle in the production of synæsthesia. It was hoped that this experiment might be included in the present investigation

but it was impossible to obtain a good case of chromæsthesia upon which to work. A newspaper appeal failed to locate a single case in the vicinity. Again we have an indication of the rarity of the phenomenon.

One other theoretical implication of the experimental findings should be mentioned. The negative nature of the results seem to indicate a limitation of the conditioned reflex-response technique, which has heretofore been considered by some (Watson, **25**) to be a tool of almost unlimited applicability. Since the results for both the group and single pairs of stimuli were negative, the data do not permit any conclusions as to the possibility of establishing a series of conditioned responses simultaneously. The results do suggest that it is not likely that the technique is going to permit a crossing-over of any two sense fields. Further experimentation, however, will be necessary before definite conclusions can be drawn in regard to either of these questions.

(Manuscript received March 2, 1933)

References

1. Allen, F., The significance of reflex action in color vision, *British Medical Journal*, 1930, No. 3642, 683–684.
2. Argelander, A., *Das Farbenhören und der Synæsthetische Faktor der Wahrnehmung*, pp. 172, Gustav Fischer, Jena, 1927.
3. Carnaz, Ch. A., *Des Abnormites congenitales des jeus et de leurs années*, Lusanne, 1848 (Quoted by Krohn).
4. Cason, H. B., The conditioned pupillary reaction, *Journal of Experimental Psychology*, 1922, **5**, 108–146.
5. Ellis, Havelock, Mescal: A study of a divine plant, *Popular Science Monthly*, 1902, **61**, 52–71.
6. Fêrè, Ch., Sur les Affets Generaux des Excitations des Organs des Sens, *Le Bulletin Médical*, 1887, Nos. 83 and 87.
7. Galton, Sir Francis, *Inquiries into Human Faculty*.
8. Gautier, T., Le club des hachichiens, *La Presse*, July 10, 1843.
9. von Goethe, J. W., Zur Farbenlehre, *Naturw. Schriften*, 4Bd., 2 Abt.
10. Hoffman, L., *Versuche einer Geschichte der Malerischen Harmonie überhaupt*, Halle, 1786 (Quoted by Krohn).
11. Kaiser, H., Assoziation der Worte mit Farben, *Archiv. f. Augenheilkunde*, 1882, **11**, 96.
12. Krohn, W. O., Pseudo-chromæsthesia or the association of colors with words, letters, and sounds, *Amer. Jour. of Psych.*, 1892, **5**, 20–41.
13. Lewes, George H., *Problems of Life and Mind*, 3rd Series, Houghton, Osgood and Co., Boston, 1880, pp. 280–285.

14. LUNDBORG H., Die Vererbung des Farbenhörens und gewisser anderer psychische Eigenschaften innerhalb einer Schwedischen Geschlechts, *Act. med. Scand.*, 1923, **59**, 221–229.
15. LUSSANA, F., *Fisologia der colori*, Pandone, 1873 (Quoted by Krohn).
16. MAHLING, F., Das Problem der Audition Colorée, *Arch. f. d. ges. Psych.*, 1926, **57**, 165–301.
17. NUEL, J., *Dictionnaire Encyclopedique des Sciences Médicales*, Vol. 83, Article on Retina.
18. NUSSBAUMER, J. A., Über Subjektiv Farbenempfindungen usw., *Med. Wochenschrift*, Nos. 1, 2, and 3, Vienna.
19. PAVLOV, I. P., *Conditioned Reflexes*, Humphrey Milford, Oxford Press, 1927, pp. 430.
20. POUCHET AND TOURNEUX, *Precis d'Histologie Humaine et d'Histogenie*, 2nd Edition, 1878.
21. SCHENKL, A., Über Assoziation der Worte mit Farben, *Prag. Med. Wochenschrift*, 1883, **10**, 94, and **11**, 101.
22. STEVENS, J. S., Colors of letters, *Popular Science Monthly*, 1892, **40**, 697.
23. TITCHENER, E. B., *A Textbook of Psychology*, Macmillan, New York, 1921.
24. WARREN, H. C., Psychology of the central nervous system, *Psych. Rev.*, 1921, **28**, 249–269.
——, The significance of neural adjustment, *Psych. Rev.*, 1922, **29**, 481–489.
25. WATSON, J. B., The place of the conditioned reflex in psychology, *Psych. Rev.*, 1916, **23**, 89–116.
26. WEISS, A. P., Behaviorism and behavior, I and II, *Psych. Rev.*, 1924, **31**, 32–50, and 118–149.
27. WHEELER, R. H., The synæsthesia of a blind subject, *Univ. of Oregon Publications*, 1920, **1**, No. 5.
28. WOODWORTH, ROBERT S., *Psychology*, Revised Edition, Henry Holt and Company, New York, 1929.

THE AFFECTIVE VALUE OF COLOR AS A FUNCTION OF HUE, TINT, AND CHROMA

BY J. P. GUILFORD

University of Nebraska

In a recent study [1] in which the writer participated, affective values of the eighteen Milton Bradley hues of medium tint (or brilliance) were obtained from 1279 subjects, including 464 men and 815 women. The colors had been presented in the form of squares $3\frac{1}{2}$ inches in size and they were judged by the method of paired comparisons. It seemed to the writer that the results were reliable enough to merit closer examination than was given to them in the previous study. Certain aspects of the data suggested that there might be some fundamental functional relationship between affective value and the three variables of color, namely, hue, tint, and chroma. It might even be possible, it was thought, to find some empirical equations by which the affective values of colors could be predicted from their known attributes. Such quantitative laws would be of considerable practical and theoretical consequence if they could be established.

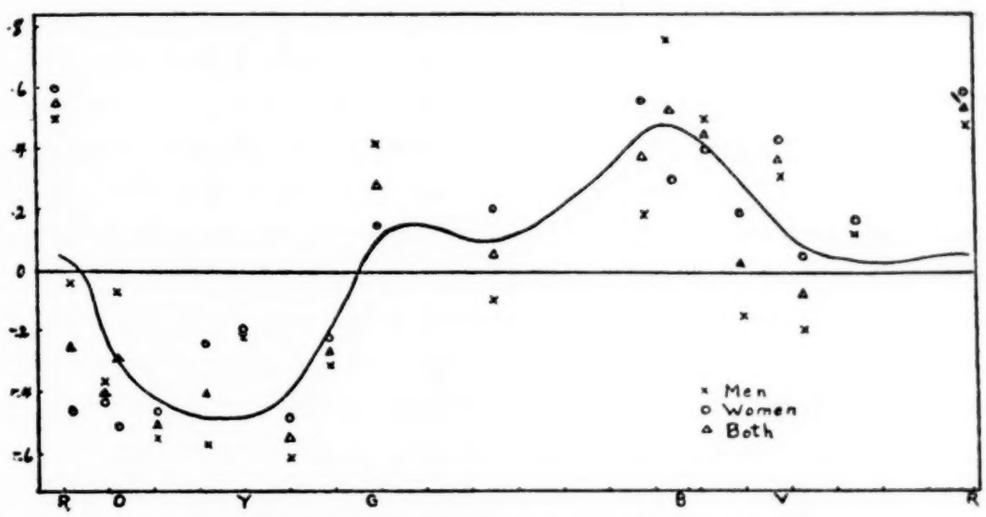
In Fig. 1 I have attempted to show graphically the relationship that appears to hold between affective value and hue. The hues were evaluated by means of the Munsell *Book of Color*. Seven different observers coöperated in rating the eighteen colors in good diffused daylight on three different occasions. Ratings were made for tint (value) and chroma as well as for hue. I have been assured by the Bureau of Standards that the Munsell color ratings are based upon equal psychological steps. I have also checked roughly the agreement between the Munsell scale of hues and the spacing of the principal colors as given by Troland,[2] and the two seem to coincide fairly well. The affective values used in Fig. 1 were

[1] W. E. Walton, Ruth B. Guilford and J. P. Guilford, Color preferences of 1279 university students, *Amer. J. Psychol.*, 1933, **45**, 322–328.

[2] L. T. Troland, in Murchison's *Foundations of Experimental Psychology*, 1929, 190.

obtained by means of the writer's method of computing scale values from paired comparisons.[3]

Although the original data seem to show constant sex differences, those differences are not extensive enough to warrant the drawing of two curves. It will be seen in Fig. 1 that the similarities between the two sexes far outweigh the differences, and that a definite trend can be made out in the scattering of the points. The curve was drawn in its present form by taking several things into consideration. First, a graphic smoothing process was resorted to. This helped to smooth out the irregularities, especially in the regions from *R* to *G*, and from *B* to *V* where the points are more numerous.



Hues (Spaced according to the Munsell color scheme)

FIG. 1. Showing the apparent functional relationship between affective value and hue as found from 815 women and 464 men.

It was pointed out in the previous article that the simple colors, *R*, *O*, *Y*, *G*, *B*, and *V*, tend to have unusually high affective values, as compared with their immediate neighbors. The resulting peaks in the distribution of points were therefore somewhat discounted. The regions between *G* and *B*, and between *V* and *R* have so few points, that much guess work has admittedly been necessary. But the final curve, with its three maxima and its three minima seem in line with the results from previous investigators and with the results

[3] J. P. Guilford, The method of paired comparisons as a psychometric method, *Psychol. Rev.*, 1928, 35, 499 f.

which the writer later obtained under better conditions. In general it has been found that *B*, *G*, and *R* are the most preferred hues, and that *V*, *O*, and *Y* are less preferred.

But before we say that the curve in Fig. 1 represents the essential relationship between affective value and hue, let us be sure that tint and chroma are not responsible for its shape to some extent. The tint and chroma of the eighteen colors are decidedly not uniform. The reds, blues, and violets are obviously darker than the oranges and yellows. The reds and blues are more saturated than the yellows and greens, and perhaps the violets and oranges. In fact the curve in Fig. 1 as it stands, approaches very nearly the reverse of a curve of brightness. The shape might possibly be due to the fact that the darker colors are preferred to the lighter ones and the more saturated ones to the less saturated.

In order to see what the relationships of affective value to tint and chroma were, correlations between affective value and the Munsell ratings were obtained. The raw correlation between affective value and tint was $-.72$ for the men and $-.62$ for the women. The correlation between affective value and chroma was .43 for the men but $-.44$ for the women. If these raw coefficients mean anything at all, men are slightly inclined to prefer saturated colors and women to prefer unsaturated ones. The partial coefficients, with chroma held constant, were even more decisive, being $-.75$ and $-.74$. The partial coefficients, with tint held constant, were .52 and $-.63$ for men and women respectively.

These coefficients might or might not represent the truth. Since there is a strong relationship between hue and tint in these eighteen colors, and since there is some relationship between hue and chroma, and since the factor of hue has not been held constant, we cannot say much about the validity of these coefficients. Later experiments have demonstrated that three of them are entirely misleading. I have presented them here to remind the reader that raw correlation coefficients, and even partials, may give erroneous results when *all* other factors have not been held constant, by experiment or by statistical treatment. It was therefore decided to secure new data from

colors in which the factors of hue, tint, and chroma were better controlled, and to use a much larger number of colors.

Brief Resumé of Some Previous Investigations.—Before entering into a discussion of the main experiments of this study, I wish to review very briefly some of the relationships previously found between affective value and the three attributes of color. It is true that with a few exceptions the interest has been in hue. It is also true that with one or two exceptions the factors of tint and chroma have not been held constant when the effect of hue has been studied. Since the different investigators have used different sets of colors, however, the same hue *may* have been presented in varying degrees of tint and saturation. It is not at all likely that all of them have used the Milton Bradley colors for their stimulus material. I shall present first some of the conclusions concerning the effect of hue, with adults, then with children, and finally with different races. Not so long ago a prominent psychologist writing for popular consumption in the *American Mercury* made the dogmatic and somewhat sweeping statement that preferences for color are merely a matter of convention.[4] This statement was not supported by even a shred of evidence, and it is made in the face of what seems to me to be overwhelming evidence to the contrary. Let the following array of evidence speak for itself.

Jastrow[5] (1897) found that *B* and *R* were the most preferred and that *Y* and *O* the least. Men preferred *B* to *R*, but women placed *R* first. Luckiesh[6] (1918) reports that the maximum affective values were in the *R* and *B*, and the minimum values were at lemon yellow and violet purple. Dorcus[7] (1926) found the order, from greatest to least affective value, to be *B P* (purple) *G R O Y*. Walton and Morrison[8] (1931), working with colored lights instead of papers and including white along with colored stimuli, found the order to be: Male, *B R G Y W;* female, *G R B Y W;* both

[4] K. Dunlap, *Amer. Mercury*, 1930, **21**, 333–336.

[5] J. Jastrow, The popular æsthetics of color, *Pop. Sci. Mon.*, 1897, 361 ff.

[6] M. Luckiesh, *The Language of Color*, 1918, 186.

[7] R. M. Dorcus, Color preferences and color associations, *Ped. Sem.*, 1926, 399–434.

[8] W. E. Walton and B. M. Morrison, A preliminary study of the affective value of colored lights, *J. Appl. Psychol.*, 1931, **15**, 294–303.

sexes combined, *B G R Y W*. Katz [9] (1931), working with insane subjects, found the order to be *B G R V Y O*.

Children of school age do not differ very much from adults in their preferences. Winch [10] (1909) found that *B* and *R* were most highly preferred, and black was least preferred. Holden and Bosse [11] (1900) found that up to about the age of eight *R* was the first choice and afterwards *B*. Katz and Breed [12] (1922) found with 2500 children the prevailing order was *B G R V O*, with a rising preference for the shorter wave lengths and a declining preference for the longer wave lengths with increasing age. Garth [13] (1924) found the following orders when 1000 children were studied: Boys, *B R G V O Y W;* girls, *B G O V R Y W*. Staples [14] (1931) studied preferences from infancy to adult levels with colors equated for tint and chroma. Up to the age of two years the order of preference remained *R Y G B*, as indicated by eye-fixation tests or by reaching tests when the colors were paired with an equivalent gray. Pre-school children, age 2.5 to 5 years, prefer colors in the order: *R G B Y*, as indicated in paired comparison tests. School children, age 7 to 12 years, prefer colors in the order: *B R G Y*. College students give the order: *B G R Y*. Thus it can be seen that yellow sinks from a high place to a low at about the age of two. Red sinks from its leading position more gradually during grade school days.

When children of other races are used as subjects, very few changes in preference are found. Garth [15] (1922) compared Indians with whites and with mixed blooded children

[9] S. E. Katz, Color preferences in the insane, *J. Abn. & Soc. Psychol.*, 1931, **26**, 203–211.

[10] W. H. Winch, Color preferences of school children, *Brit. J. Psychol.*, 1909, **3**, 42–65.

[11] W. A. Holden and K. K. Bosse, The order of development of color perception and color preference in the child, *Arch. of Opthal.*, 1900, **29**, 261–278.

[12] S. A. Katz and F. S. Breed, The color preferences of children, *J. Appl. Psychol.*, 1922, **6**, 255–266.

[13] T. R. Garth, A color preference scale for 1000 white children, *J. Exper. Psychol.*, 1924, **7**, 233–241.

[14] R. Staples, The responses of infants to colors, *J. Exper. Psychol.*, 1932, **15**, 119–141.

[15] T. R. Garth, The color preferences of 559 full blood Indians, *J. Exper. Psychol.*, 1922, **5**, 392–418.

with the following orders: Full blood Indians, *R B V G O Y W;* mixed blood, *B R V W G O Y;* white, *B G R V O Y W.* Hirahashi [16] (1926) used Japanese children and teachers, finding that *B* and *G* were high in the list and *R* and *Y* low. Shikiba [17] (1927), with deranged adults and delinquent Japanese boys, found the order of preference to be: *B R V G Y O.* Hurlock [18] (1927) found that both white and negro children in choosing color names prefer blue and pink. The selection of pink is significant because it shows that red, even though unsaturated, is still one of the preferred colors. Garth and Collado [19] (1929), with Filipino children, found the order to be: *R G B V O W Y.* Education was found to raise B and W slightly, to lower *V* and *Y*, but to leave *R* unchanged as the highest preference. Garth, Ikeda and Langdon [20] (1931) came to the conclusion that the Japanese children prefer essentially the same colors as the white, Indian, and negro children do.

No claim is made that the above list of investigations is complete. The number of examples is amply sufficient, however, to show that in spite of differences in experimenter, color material, method of measurement, age, race and sex of the subjects, the order of preferences for the different hues is a rather uniform phenomenon. I am not prepared to maintain that conventions have nothing to do with color preferences, but certainly whoever asserts that color preference is *merely* a matter of convention is flying in the face of the facts. The differences that do occur, might be attributed to differences in brightness or chroma, in the size, shape, and background of the color stimuli, and to some extent to the age and education of the subjects. Most of these investigations used only a

[16] B. Hirahashi, Some experiments on beauty of color, *Jap. J. Psychol.*, 1926, 1, 406–432.

[17] T. Shikiba, Color preferences of deranged persons and delinquent boys, *Jap. J. Psychol.*, 1927, 2, 677–700.

[18] E. B. Hurlock, Color preferences of white and negro children, *J. Comp. Psychol.*, 1927, 7, 389–404.

[19] T. R. Garth and I. R. Collado, Color preferences of Filipino children, *J. Comp. Psychol.*, 1929, 9, 397–404.

[20] T. R. Garth, K. Ikeda, and R. M. Langdon, The color preferences of Japanese children, *J. Soc. Psychol.*, 1931, 2, 397–402.

small number of colors, even as few as five, but most of them used an adequate number of subjects.

Of the investigations in which the factors of tint and chroma were at all controlled, very little new is learned about preferences for hue. Walton and Morrison [21] found that when colored lights are equated for brightness by the flicker method, the order of preference remains about the same except that *R* seems to gain. The new orders were: For men, *R B G Y W;* for women, *G R B Y W;* for both combined, *R B G Y W*. Chroma was not equated but *R* and *Y*, being reduced in brightness, lost in chroma, whereas *B* lost none. Dorcus,[22] when equating both tint and chroma, and in so doing reducing the chroma of his colored papers, found that *R* was *less* pleasing than before, which is in direct opposition to the results of Walton and Morrison. *Y* also was less pleasing, and purple gained in affective value. It is difficult to draw conclusions from these limited and somewhat conflicting results.

The effect of tint upon affective value is suggested in only five previous studies. Jastrow [23] found that the darker colors were preferred. But judging from the results represented in Fig. 1, where the darker colors also were preferred, we can only say that the effect of hue may be responsible for this apparent effect of tint. Major [24] (1895) concluded that tint has little if any effect, and Dorcus [25] came to a similar conclusion. Miss Washburn [26] (1911), however, found that when using eighteen tints and eighteen shades of the Milton-Bradley series, the lighter colors were preferred to the darker. Luckiesh [27] (1923) agrees with Miss Washburn, and the results of the present investigation will concur in this, that the lighter the tint the greater is the affective value, other things being equal.

[21] W. E. Walton and B. M. Morrison, *Op. cit.*, 300.

[22] R. M. Dorcus, *Op. cit.*

[23] J. Jastrow, *Op. cit.*

[24] D. R. Major, On the affective tone of simple sense impressions, *Amer. J. Psychol.*, 1895, **7**, 57–77.

[25] R. M. Dorcus, *Op. cit.*

[26] M. F. Washburn, A note on the affective value of color, *Amer. J. Psychol.*, 1911, **22**, 114.

[27] M. Luckiesh, *Light and Color in Advertizing and Merchandizing*, 1923, 33 ff.

The effect of saturation has received much more attention than the effect of tint. Cohn [28] (1894) was one of the first to point out that saturated colors are preferred to unsaturated ones. Bradford [29] (1913) and Luckiesh [30] (1918) agree with Cohn in this. Major,[31] however, obtained no decisive results, but was inclined toward the conclusion that the unsaturated colors are the more preferred. Titchener [32] pointed out that there are two types of subjects. Type I prefers saturated and type II unsaturated colors. Type I is probably the prevailing one. Miss Washburn [33] contributed the interesting fact that saturated colors are preferred in small areas and unsaturated ones in large areas. The factor of size may be so important that conflicting results could be traced to this principle.

The preference of colored stimuli over those entirely uncolored is probably universal, since black and white are usually at the bottom of the list as in the results of Garth and others cited above. Staples [34] found that even among infants of a few months a colored disc is fixated longer than a colorless one, and at later ages reaching tests and paired comparisons indicate preferences for colored versus non-colored stimuli. But Dorcus,[35] who used the same hues in saturated and in unsaturated form, found little difference. It can be said in criticism of his results, however, that he used only six hues and that the difference in saturation was very slight. It ranged from six to ten points on the Munsell scale for the 'saturated' colors, and all the 'unsaturated' colors had a rating of five points. No final conclusions about the effect of differences in chroma can be drawn from such limited data as those just cited.

[28] J. Cohn, Experimentelle Untersuchungen über die Gefuhlsbetonung der Farben, Helligkeiten und ihrer Combinationen, *Phil. Stud.*, 1894, **10**, 585 f.

[29] E. J. G. Bradford, On the æsthetic value of the perceptive types in color appreciation, *Amer. J. Psychol.*, 1913, **24**, 245 ff.

[30] M. Luckiesh, *Op. cit.*, 1918, 180.

[31] D. R. Major, *Op. cit.*

[32] E. B. Titchener, *Experimental Psychology*, 1901, 1, 2, 151 ff.

[33] M. F. Washburn, *Op. cit.*

[34] R. Staples, *Op. cit.*, 126.

[35] R. M. Dorcus, *Op. cit.*

Plan of the Present Investigation

Selection of the Color Stimuli.—Several considerations were used as guides in the selection of stimuli. First of all, it was desirable to have if possible eighteen hues, with tint and chroma held approximately constant. Secondly, it was important to have more than one sample of every hue. Thirdly, a wide variety of tints and chromas was needed to give those two variables a full opportunity to show their effects. Fourthly, the number of colors could not be enlarged without limit, especially if the method of paired comparisons were to be used. The material had to be selected from the Milton Bradley series of ninety colors because of the prohibitive cost of large samples of better standardized colors like the Munsell series. It would not have been wise to use the entire set of ninety colors, for taking the series as a whole, even if there are five tints for every hue, all the yellows would have an average brightness value higher than any five reds or blues, and the five blues would have had a lower average brightness than most colors of other hues. In other words, the correspondence between hue and brightness would have been approximately the same as for the middle row of eighteen hues which were used in the study represented in Fig. 1. This was one of the errors that had to be avoided.

As a preliminary aid, the writer made some approximate ratings of the ninety colors with the use of the Munsell *Book of Color*. Next, eighteen different hues were selected, in colors having approximately equal tint and chroma. As later more careful ratings proved, the ranges of tint and chroma were not as small as one would like them to be, but they are probably the smallest ranges obtainable from the Milton Bradley series. The range in tints, as can be seen in the first eighteen colors in Table II, is from 4.3 to 7.8, with most of the values 5 or 6. The range of chromas was from 6.8 to 12.4, with the majority of them close to 9 or 10. Taking the tints as a whole, even though not exactly equal, they are distributed along the scale of hues in a very unsystematic manner. The chromas, however, are still high in the red and low in the blues and greens. This could not be avoided in selecting the first eighteen colors. This defect was corrected to some extent in making the remaining selections by using poorly saturated reds and more highly saturated blue-greens.

Next, eleven colors were selected, scattering over the whole range of hues, and chosen so as to give a great range of brightness values. These colors would of course be generally low in saturation. They ranged in tint from 3.1 to 8.9. They ranged in chroma from 4.1 to 8.4. A final group of eleven colors was chosen, varying as much as possible in hue, and consisting of only very saturated colors. The range in tint was from 2.9 to 8.2. The range in chroma was 8.5 to 13.6. Table I presents the forty colors, with the Milton-Bradley names in the second column. The sub-scripts which vary from 1 to 5, indicate the tint, 1 being the lightest for each hue, and 5 the darkest. The next three columns, headed H, T, and C, are the evaluations obtained from the Munsell *Book of Color*. The manner in which these values were obtained will be described further on.

Securing the Ratings of Affective Value.[36]—Having forty colors, the method of paired comparisons was out of the question. This method would have required 780 pairs to be judged. It was decided to use a rating scale method instead. For a double

[36] The writer wishes to express here his indebtedness to Miss Elysbeth C. Allen, Wolfe Memorial Fellow, who conducted all the experiments, and to the subjects, among whom were Professors D. A. Worcester, S. M. Corey, and Dr. D. W. Dysinger.

scale like the *PU* continuum, it has been found that a nine point scale is best,[37] and this kind of a scale was used. It had already been found to work well in judging the affective values of odors.[38] The rating scale method has an additional advantage over the method of paired comparisons in that the indifference point of the series of colors can be located.

The stimulus colors were in the form of 4 inch squares, pasted upon gray photo mount cardboard with a brightness value of exactly 5 on the Munsell scale. Over a laboratory table 30″ by 54″ was built an enclosure of the same gray cardboard, extending about 30″ high. The subject sat at a headrest at one end of the table and the colors were exposed in a small window at the other. The room was darkened. The colors were illuminated by the direct lighting of four 50 w bulbs, two on either side of the subject but shielded from him. All bulbs were frosted, one of them on either side being white and the other a bluish Mazda 'daylight' lamp. The combination seemed to give a uniform mixture that resembled daylight very closely. The window in which the colors were exposed was shielded by a shutter of the same gray material as the background while *E* made the changes of stimulus cards. *E* gave a 'Ready' signal, tripped the shutter with a foot lever, and exposed the color for 5 sec.

The instructions for *S* were always at hand in printed form and they read as follows: "After a preliminary 'Ready' signal, the stimulus card will be exposed for 5 seconds. You are to judge the pleasantness of the color presented, being particularly careful to judge it as a color. Your report of pleasantness or unpleasantness is to be given in terms of the following nine point scale:

9. greatest possible pleasure
8. extremely pleasant
7. moderately pleasant
6. mildly pleasant
5. indifferent, neither *P* nor *U*
4. mildly unpleasant
3. moderately unpleasant
2. extremely unpleasant
1. greatest possible unpleasantness."

More than forty colors might well have been used, but during the same hour of observations the subject was also called upon to judge forty-five combinations of colors. The combinations contained some of the single colors among the forty. On one day the single colors would be judged during the first part of the sitting, and on the next occasion the combinations would be judged first. The results from the color combinations will form the basis of another report. The judgment of forty single colors and forty-five combinations during the same sitting did not appear to fatigue the subjects, especially since the change from the one set to the other brought some degree of novelty into the situation, and since rest pauses were inserted. The order of presentation of the colors was determined partly by chance. The cards were shuffled, and then the order was examined to see that no hue followed itself, that no decided contrasts in tint occurred and as nearly as preliminary guesses could determine, that no decided contrasts in affective value might occur. On alternate days the order of presentation was exactly reversed.

---

[37] This has been shown by Professor E. S. Conklin in the *Univ. of Oregon Pub.*, 1923, **2**, pp. 26.

[38] W. Spence and J. P. Guilford, The affective value of combinations of odors *Amer. J. Psychol.*, 1933, **45**, 495–501.

Ten subjects took part in the experiments, five men and five women. It was thought more desirable in using the scale of values method to secure several repeated judgments from a few subjects than to secure only one set of judgments from each of a large number of judges. It sometimes takes one or two sets of judgments for *S* to become accustomed to the use of the rating scale. It has been found by other investigators [39] that the early trials in judging colors are more largely controlled by associations aroused by the colors, and that in later trials the associations tend to disappear.

Seven of the ten subjects were trained to some extent in laboratory procedures, three of them being instructors in psychology, and four of them being graduate students. The remaining three were undergraduates. None of them were trained in the fine arts. Each one of the ten judged the set of colors five times on five different days covering a period of about one month. The reliability of their ratings will be discussed shortly

The Ratings of Hue, Tint, and Chroma.—After the judgments of affective value were completed, the ratings of the forty colors for hue, tint, and chroma were obtained from the same ten subjects. The colors were exposed under the same illumination as before and the Munsell Book of Color was placed beside them. *S* selected first the nearest hue in the Book, or interpolated whenever necessary, then judged the tint, and the chroma. Each *S* judged all forty colors on two different days, and the ratings given in Table II are the means of twenty such judgments.

One might, in a study like this, attempt to use in place of the Munsell ratings for hue, the predominating wave length of the various colors. To find a functional relationship between affective value and wave length would introduce into psychology a new and interesting psychophysical law. But this procedure faces one very serious difficulty. The psychological series of hues is a complete circle, including within it the purples for which there are no assignable physical wave lengths. We are therefore thrown back upon the plan of finding a purely psychological law, of the relationship between experienced hues, measured in equal phenomenal steps if possible, and the affective values that go along with them.

Reliability and Validity of the Ratings of Affective Value.—One could now, of course, treat the affective values for each *S* separately, noting the effect of hue, tint, and chroma upon his individual ratings. This would demonstrate whether or not the same functional relationship, or at least the same type of relationship, holds for all ten of the *S*s. But in this investiga-

[39] J. Cohn, *Op. cit.*, 596.

tion, as in all others on color preference, we are interested in the generalized reactions of a large group of individuals. We therefore deal in averages.

This being true, the question arises, how nearly will ten *Ss* chosen at random give results that tally with the generalized laws of affective value? How representative are the ten *Ss* used in this study? Furthermore, since we are accustomed to treat the sexes separately in problems on color preference, it means that in this study we have to let five individuals speak for their own respective sex group. In order to derive some notion of the universality of the affective values with which we are dealing here, the following procedure was resorted to.

TABLE I

INTERCORRELATIONS OF THE INDIVIDUAL JUDGMENTS

	Women						Men				
	A	*B*	*C*	*D*	*E*		*F*	*G*	*H*	*I*	*J*
B	.54					*G*	.06				
C	.40	.82				*H*	.39	.20			
D	.34	.63	.47			*I*	.38	.67	.54		
E	.55	.73	.46	.55		*J*	.41	.03	.56	.30	
r_{ag}	.58	.995	.71	.65	.78		.48	.29	.75	.64	.50
r^2_{ag}	.34	.99	.50	.42	.60		.23	.08	.56	.41	.25
w_a	.88	99.5	1.42	1.13	1.94		.63	.32	1.70	1.08	.67

Mean intercorrelation	.548	.354
Theoretical $\overline{\text{p.e.}}_{\text{tetrad}}$	.055	.058
Actual $\overline{\text{p.e.}}_{\text{tetrad}}$	.072	.117
r_{pg}	.995	.855

In each group of five *Ss*, the average ratings of each *S* were correlated with the average ratings of every other *S*. These intercorrelations are presented in Table I. For the women, the coefficients range from .34 to .82, with a mean of .548. For the men they are not so high, but they are all positive, ranging from .06 to .67, with a mean of .354. It was assumed, next, that these tables of intercorrelations represented a two-factor pattern of the Spearman type. This would mean that any *S*'s judgments are determined by the *g*-factor (or factors), those determiners that all the *Ss* hold in common, and in addition they are determined by a set

of factors peculiar to every *S* alone. The latter can be called *s*-factors.

The tetrad differences [40] were found for both sets of intercorrelations, and their mean p.e.s were found to be not significantly larger than would be expected by chance. The two-factor assumption may therefore be made. Then the correlations of each *S*'s ratings with the *g*-factor were computed. These appear in Table I. They range from .58 to .995 for the women and from .29 to .75 for the men. It is interesting to notice that one of the Ss, B, had a correlation with the *g*-factor that is perfect. The squares of the r_{ag} values are also given, since they measure the percent of the judgments in each case that can be said to be determined by the *g*-factor. One of the men was only 8 percent determined by the *g*-factor, but one of the women was 99 percent so determined. It is significant, perhaps, that three *S*s with some of the lowest correlations with the *g*-factor, *S*s *D*, *G* and *J*, were the undergraduate, less experienced judges.

The weights that should be assigned to each *S*'s judgments are given under the label of W_a. Except for the weight of subject B, they are of about the same order of magnitude. If in pooling the ratings of the five *S*s in each group, standard measures were used instead of the original judged values, and if those standard measures were weighted according to W_a of Table I, a highly valid final measurement of the *g*-factor would be obtained, especially for the women. The correlations for such weighted pools with the *g*-factor would be .995 and .855 for the women and men respectively. The unweighted pools, which were actually used, would give measures of the *g*-factor, of course, but probably with a somewhat lower degree of validity. I have spoken of *the g*-factor, as if there were only one, and that that one appears in both sets of data. The proof of this requires that the Spearman tetrad difference test be applied to a single combined table of intercorrelations from all ten *S*s. This has not been done, since it is not entirely relevant to the tests of validity which are being made here. But I am inclined to believe that the

[40] Spearman's procedures were followed as given in K. J. Holzinger's *Statistical Resume of the Spearman Two-Factor Theory*, 1930, University of Chicago Press.

subsequent results of this report will reveal such striking similarities between the two sex groups that it would be safe in saying that the *g*-factors are approximately identical.

The average affective values for the women and for the men are to be found in Table II, in the columns labelled AV. It is around these that the following discussion revolves.

TABLE II

EVALUATIONS OF HUE, TINT AND CHROMA, AND OF AFFECTIVE VALUE FOR THE FORTY COLORS

Color No.	M-B Name	*H*	*T*	*C*	Women			Men		
					AV	*z*	*z'*	AV	*z*	*z'*
1	*R*2	0.0	4.6	12.4	4.76	− .69	− .95	5.46	− .04	− .71
2	*OR*2	0.0	6.1	10.6	5.56	.11	− .34	6.24	.74	.13
3	*RO*4	.8	4.6	11.4	5.96	.61	.08	5.78	.51	.00
4	*O*4	1.0	5.1	10.6	5.24	− .08	− .29	5.52	.37	− .09
5	*YO*4	2.0	5.6	10.2	4.20	− .20	− .50	3.86	− .42	− .88
6	*OY*5	3.0	5.5	8.9	3.12	− .28	− .44	3.00	− .25	− .49
7	*Y*4	3.9	7.8	9.9	4.44	1.22	.43	4.32	1.26	.51
8	*GY*4	5.0	6.3	7.5	4.36	.43	.21	3.44	− .26	− .39
9	*YG*4	6.1	4.3	7.3	5.92	.65	.92	5.18	.38	.58
10	*G*4	7.0	6.2	9.5	5.76	− .12	− .35	5.76	.30	− .18
11	*BG*3	9.7	5.7	6.8	5.24	− .23	− .23	4.84	− .22	− .14
12	*GB*3	13.4	5.0	9.1	7.76	− .12	− .12	6.44	.02	− .17
13	*B*2	13.9	5.5	10.7	8.04	.04	− .27	6.54	.07	− .46
14	*VB*2	15.0	5.7	9.3	7.56	.01	− .23	6.24	.07	− .26
15	*BV*3	15.0	4.9	10.2	7.44	− .11	− .25	6.04	− .16	− .52
16	*V*2	16.0	6.2	7.4	6.80	.00	− .18	5.20	− .58	− .68
17	*RV*3	16.1	4.7	10.4	6.72	− .08	− .19	5.02	.27	− .09
18	*VR*2	17.8	5.6	10.3	5.04	− .28	− .53	5.16	− .09	− .58
19	*OR*1	0.1	7.5	7.3	5.84	.41	− .07	5.62	.12	− .16
20	*RO*5	1.0	3.8	8.4	5.44	.12	.41	5.82	.62	.72
21	*O*1	1.2	6.9	6.3	5.64	.39	.14	5.54	.49	.47
22	*OY*1	3.1	8.9	5.9	4.68	1.33	.65	4.46	1.30	1.04
23	*GY*5	4.8	5.4	7.4	3.08	− .74	− .73	2.90	− .56	− .53
24	*G*1	7.1	8.4	4.8	5.60	− .30	− .76	4.80	− .70	− .71
25	*GB*5	13.7	3.7	5.5	6.44	− .51	− .92	4.86	− .64	− 1.05
26	*B*1	13.8	7.4	7.1	7.60	− .40	− .84	5.90	− .58	− .81
27	*BV*1	15.1	7.0	6.6	6.88	− .62	− .91	5.60	− .55	− .84
28	*RV*1	16.6	8.2	4.1	5.96	− .16	− .51	4.96	− .61	− .47
29	*VR*5	18.6	3.1	6.2	4.48	− .69	− .03	5.22	− .04	.52
30	*R*5	0.0	2.9	8.5	5.28	− .17	.33	6.50	1.00	1.22
31	*OR*3	.2	5.0	13.6	4.24	− 1.20	− 1.68	5.34	− .14	− 1.07
32	*O*3	1.2	5.9	12.2	5.28	.03	− .52	6.02	.95	.11
33	*YO*3	2.0	6.9	11.9	5.64	1.24	.47	5.82	1.52	.58
34	*Y*3	4.1	8.1	10.6	5.52	2.27	1.37	6.30	3.22	2.30
35	*GY*3	5.0	8.2	8.8	5.60	1.67	.88	6.24	2.62	1.99
36	*YG*3	6.0	8.2	8.6	6.52	1.28	.51	5.38	.68	.08
37	*G*3	6.9	6.8	9.4	6.32	.47	− .04	5.94	.47	− .05
38	*B*4	14.3	3.7	12.2	7.52	− .41	− .44	6.98	.58	.07
39	*V*3	15.6	4.8	9.8	6.52	− .66	.76	5.94	− .01	− .34
40	*VR*3	18.1	3.4	9.4	6.32	1.12	1.41	6.70	1.48	1.48

Taking these as representative, if not universal averages of affective values of the forty colors, how do they depend upon hue, tint, and chroma?

The Results

The Influence of Hue.—The influence of hue is best shown by the first eighteen colors in Table II. In these the differences in tint and chroma were as small as possible, and the distributions of tint and chroma were almost totally unrelated to the distribution of hue. The AVs for these eighteen colors

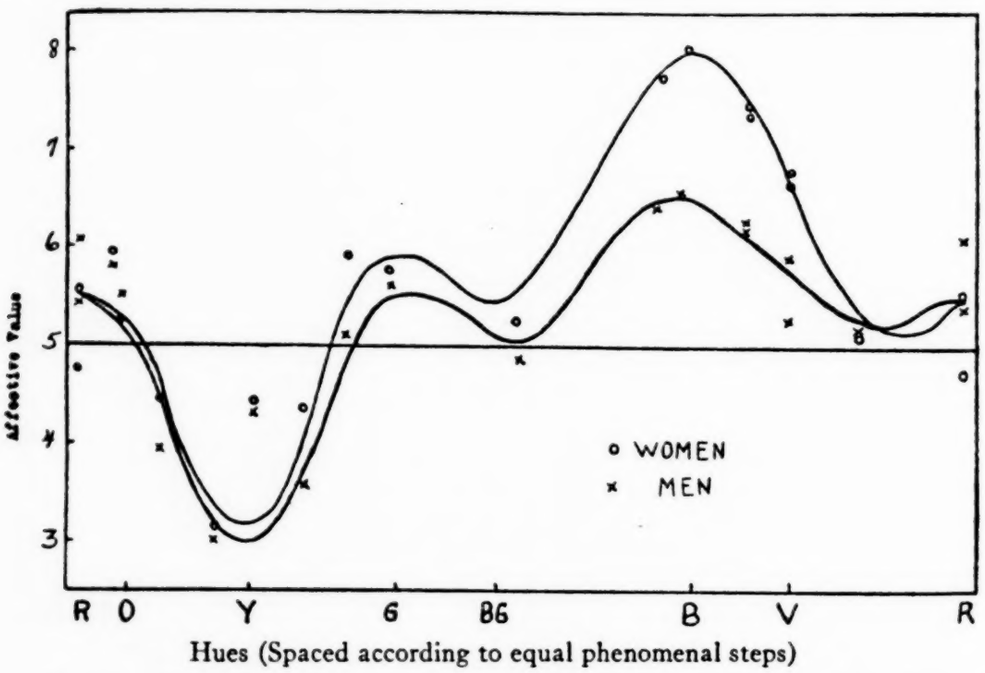


Fig. 2. Curves showing the functional relationship between affective value and hue for both men and women. The curves as drawn are based upon the first four harmonics.

are plotted against the hue positions in Fig. 2. The curves were smoothed in the same manner as those in Fig. 1. The similarity in form between these curves and the curve in Fig. 1 is so complete that one is forced to the conclusion that the AVs represented in Fig. 1 were actually determined to a very large extent by hue and to a smaller extent, if any, by tint and chroma. In other words, *R* and *B* are not most preferred because they are darker, and *Y* and *O* least preferred because they are lighter. Neither, probably, are the AVs of *R* and *B*

high merely because of a high degree of chroma. And since the curves in Fig. 2, obtained from only five *Ss*, agree in form with that in Fig. 1, obtained from hundreds of *Ss*, we have some additional evidence of the generality of the AVs used in this study, even though they were based upon a limited number of *Ss*.

The whole character of the curves, either in Fig. 1 or in Fig. 2, suggests that we have there some periodic functions. All three curves have three maxima, and three minima, and those in about the same regions. The abscissa represents a complete cycle of steps; it represents the color circle. The curves also return upon themselves, the two ends meeting at the same value at red. The curves in Fig. 2, especially, resemble closely trigonometric functions of the type that are used to represent sound waves, alternating currents, and similar phenomena.

According to the Fourier theorem, any periodic curve can be analyzed into a number of harmonics, in the form of simple sine and cosine functions. From all appearances, the curves under discussion would seem to be of a simple type, analyzable into a very few harmonics. Accordingly, the curves in Fig. 2 were drawn to scale and subjected to a Fourier analysis by means of a mechanical harmonic analyzer.[41]

The results of this analysis are presented in Table III. The process was carried no further than the first six harmonics, since the last two or three of these had rather insignificant weights. Each succeeding harmonic, no matter how small its amplitude, adds a closer approximation to the curves as they were drawn. But since the curves are to some extent a matter of guess-work, only the grosser aspects of them are worth the trouble of fitting. For this reason, too, the process was not carried beyond the sixth harmonic.

It is interesting, but to be expected from Fig. 2, that the men's and women's curves gave such similar weights or coefficients for the different harmonics. It is interesting, also, that the first and third harmonics are by far the most signifi-

[41] I am very much indebted to Major Allen P. Cowgill, instructor in mathematics, for performing the harmonic analyses and to the department of mathematics of the University of Nebraska for the use of the harmonic analyzer.

cant ones. The latter fact takes on added significance when viewed in connection with theories of color vision. This point will engage our attention later.

TABLE III

WEIGHTS OR COEFFICIENTS FOR THE FIRST SIX HARMONICS IN THE PERIODIC CURVES OF FIG. 2

Harmonic	Women			Men		
	Sine	Cosine	Amplitude	Sine	Cosine	Amplitude
1st	−1.41	−.75	1.60	−1.02	−.38	1.09
2nd	− .02	−.16	.16	− .20	.17	.26
3rd	.35	.69	.77	.23	.61	.65
4th	.22	−.03	.22	.15	−.07	.17
5th	.01	−.01	.02	− .02	.05	.05
6th	.08	.04	.09	.07	.04	.08
Constant	.67			.17		

The amplitudes give a fair idea of the relative strengths of the various harmonics. Each one, of course, gives a range of AVs twice as great as the value given in Table III. For example, the first harmonic for the women gives a range of 1.60 units above or below the mean, or a total range of 3.20. The third harmonic, likewise, gives a range of 1.54. The two, summated, give a range of 4.74. The two amplitudes may be summated in this particular case, since the one harmonic has just twice the frequency of the other so that their crests often coincide. The actual total range of AVs for the women was 4.96, so it can be seen that the first and third harmonics taken together account for most of that spread of values. The same harmonics account for 3.48 units in a total actual range of 3.98 units for the men. In order to secure maximum predictive value from these sine-cosine functions, I have used all coefficients greater than ±.10 in Table II. This includes most of the weights for the first four harmonics. The empirical equations may be stated as follows:

$$\text{For women: } AV = 5.67 - 1.41 \sin 18H + .35 \sin 54H + .22 \sin 72H - .75 \cos 18H - .16 \cos 36H + .69 \cos 54H.$$

For men: $AV = 5.17 - 1.02 \sin 18H - .20 \sin 36H + .23 \sin 54H + .15 \sin 72H - .38 \cos 18H + .17 \cos 36H + .61 \cos 54H$.

Before these equations are taken too seriously as the universal laws of preferences for hues, let it be repeated that they represent the smoothed function that was fitted to the data graphically. It may be that in drawing the curves some harmonics of higher order have been smoothed over. It may just as well be that some harmonics other than the first and third have been erroneously introduced. After much deliberation and working over the data, the writer is convinced that the first and third harmonics are correctly represented and that they are by far the most important ones. The most uncertain part of the curves is in the region between green and blue. Unfortunately the Milton Bradley colors do not represent that region sufficiently. It will take many more than eighteen hues to make sure that any harmonics of higher order have not been overlooked. There is some uncertainty, for example, as to whether yellow should be given a minor peak. The unusually high position of yellow as compared with the trend of the curves may be due to the fact that it is a simple, unmixed hue. On the other hand, all affective values which we have for different samples of yellow, are higher than their tint and chroma would warrant. The curves as presented must be taken as partly hypothetical and as partly empirical fact.

Leaving aside for the moment all theoretical implications that may follow from these periodic functions, let us see how well the AVs could have been predicted from the empirical equations alone. The forty predictions, for both men and women, were made by substituting in the above formulas the necessary values of H. The values of H are obtained from the Munsell scale with its twenty equal steps. The value of 0 was assigned to red and the rest were numbered accordingly. Next the discrepancies between the predictions and the actual AVs were computed. The discrepancy in each case is equal to the observed AV minus the predicted AV. These discrepancies are given in Table II in the columns labelled 'z.'

They are relatively small for the first eighteen colors, somewhat larger for the second group of eleven, showing that differences in brightness have their effect, and still larger for the last eleven, showing that chroma also has its effect.

Coefficients of correlation were computed between the actual and the predicted AVs. The coefficients are .84 and .55 for the women and men, respectively. When corrected for the number of variables, they are .82 and .40.[42] In each case ρ^2 is .666 and .156. The latter values are sometimes called *indices of determination*,[43] and they are said to measure the percent of the total determination of the dependent variable, AV in this case, that is due to the independent variable, in this case, hue. In other words, the AVs of the women are determined by hue to the extent of about 67 percent and by all other factors combined to the extent of 33 percent. The men's AVs are determined only about 16 percent by hue, and 84 percent by something other than hue. How much of the remaining percentages are due to tint and chroma?

The Effect of Tint and Chroma.—The determination of the AVs by tint and chroma can best be found from the discrepancies, z, for in these discrepancies the probable effect of hue has been eliminated. The regressions of z on tint and chroma seem to be linear, judging from the scatter diagrams. The correlations between these discrepancies and tint were approximately .41 and .16 for the women and men respectively. The correlations with chroma were .12 and .32 respectively. From the raw coefficients it would appear that both men and women prefer the lighter colors to the darker, and that this tendency is stronger for women than for men, and that both prefer the more saturated to the less saturated, the effect of chroma being more strongly felt by the men. This interpretation agrees with the usual result in most respects. But as it happened, there was a small amount of correlation between tint and chroma in the forty colors ($r_{TC} = -.28$), so that the *net* effect of tint and chroma cannot be seen in these raw coefficients. The *partial* coefficients between tint and chroma on the one hand and the discrepan-

[42] M. Ezekiel, *Methods of Correlation Analysis*, 1930, 121 f.

[43] *Ibid.*, 120.

cies on the other give a more truthful picture of the sole effect of tint or of chroma. These partial coefficients, corrected for the number of variables, may be found in Table IV. The

TABLE IV

PARTIAL AND MULTIPLE COEFFICIENTS AND INDICES OF CORRELATION BETWEEN AFFECTIVE VALUES AND THE THREE ATTRIBUTES OF COLOR, WITH INDICES OF DETERMINATION

	$\rho_{AH.TC}$	$\bar{\rho}_{AH.TC}$	$\bar{\rho}^2_{AH.TC}$	$r_{AT.HC}$	$\bar{r}_{AT.HC}$	$\bar{r}^2_{AT.HC}$	$r_{AC.HT}$	$\bar{r}_{AC.HT}$
Women....	.844	.816	.666	.472	.450	.203	.265	.219
Men.....	.553	.395	.156	.279	.230	.053	.388	.361

	$\bar{r}^2_{AC.HT}$	$R_{ATC.H}$	$\bar{R}_{ATC.H}$	$\bar{R}^2_{ATC.H}$	$P_{A.HTC}$	$\bar{P}_{A.HTC}$	$\bar{P}^2_{A.HTC}$
Women....	.048	.473	.425	.181	.880	.844	.712
Men.....	.130	.417	.359	.129	.652	.508	.258

net correlations with tint are .47 and .28. The net correlations with chroma are .27 and .39. The percentages of determination of the AVs by tint and chroma are given in Table IV. We find that tint determines the AV about 20 percent for the women and 5 percent for the men. Chroma determines the AV to the extent of about 5 percent for the women and 13 percent for the men. From this it will be seen that chroma is almost as strong a factor as hue for the men.

The combined effect of tint and chroma is given by the partial coefficients $R_{ATC.H}$ as presented in Table IV. They are .47 and .42 for women and men respectively. Corrected and squared, they show that tint and chroma combined determine the AV to the extent of 18 percent for the women and 13 percent for the men. Comparing these coefficients of determination with those for hue, we find that for women, the effect of hue is more than three times as strong as the combined effect of tint and chroma, and for men, tint and chroma combined have almost as much effect as hue. The computed relative effects of hue, tint, and chroma will depend to some extent upon the degree of perfection of the 'fit' of the curves for hue, and whether the effect of hue has been successfully eliminated. The above percentages of deter-

mination are given merely as approximations under the conditions of this study. They are probably within reach of the truth, and they will at least show what might be done in finding an accurate evaluation of the factors determining affective value.

Regression equations, predicting the discrepancies (z) from tint and chroma were established. They are as follows:

$$\text{For women: } z = .24T + .09C - 1.97.$$
$$\text{For men: } z = .15T + .16C - 2.05.$$

We may now combine these equations with those for hue, giving a complete quantitative law by which affective value can be predicted from all three attributes. This is possible because

$$z = AV - P \text{ (P being the prediction from hue alone)}$$

or,

$$\text{AV} = z + P.$$

But

$$P = f(H), \text{ and } z = f(TC).$$

Therefore,

$$\text{AV} = f(H) + f(TC).$$

The complete equations are:

For women: $\text{AV} = 3.70 + .24T + .09C - 1.41 \sin 18H + .35 \sin 54H + .22 \sin 72H - .75 \cos 18H - .16 \cos 36H + .69 \cos 54H.$

For men: $\text{AV} = 3.12 + .15T + .16C - 1.02 \sin 18H - .20 \sin 36H + .23 \sin 54H + .15 \sin 72H - .38 \cos 18H + .17 \cos 36H + .61 \cos .54H.$

Predictions of AV were made with these complete formulas. The discrepancies between these predictions and the observed AVs were obtained, and they appear in Table II in the columns labelled 'z'.' It can be seen that these discrepancies are usually smaller than the former ones, especially among the last

twenty-two colors where the effects of tint and chroma are more potent. The multiple correlations between the predictions from all three attributes were obtained and they will be found in Table IV. Uncorrected, they are .88 and .65 for women and men respectively. Corrected and squared, they show that in hue, tint, and chroma we have 71 percent of the factors which determine the AVs for the women, but only 26 percent of the factors controlling the preferences of the men.

These percentages of determination do not check so very closely with the sums of percentages of determination from hue, tint, and chroma taken separately. From Table IV we find for the women that the percentages of determination are 67, 20, and 5, for hue, tint, and chroma, respectively. The sum is 92, as compared with 71 percent which was found from the multiple coefficient of correlation. For the men, the three percentages are 16, 5, and 13, with a sum of 34, as compared with 26 percent from the multiple correlation. I am not entirely prepared to account for these discrepancies. It would seem that any statistical device which pretends to measure the percentage of determination that can be attributed to different factors, when properly applied, ought to yield a better internal consistency than this. One cause of the lack of consistency here may be in the more drastic corrections that are applied to the multiple coefficients. The larger the number of variables in the equation used for prediction, the greater is the correction. The number of variables in the complete equations here becomes rather large. And it is not certain just how many 'variables' and hence how many 'degrees of freedom' there are in sine-cosine functions. In spite of these uncertainties, however, the writer believes that the obtained percentages of determination give a fair idea of the relative importance of hue, tint, and chroma for affective value.

Discussion

Evaluation of the Results.—From a practical standpoint, the prediction of affective value from the three attributes of color, as in the complete formulas given above, still leaves much to be desired. The formulas contain entirely too many terms to be

readily applied. And yet the curves, smoothed graphically, have been made about as simple as the data justify. In addition, the lack of determination of affective value by hue, tint, and chroma, especially in the case of the men, makes predictions of insufficient accuracy. And yet, the percent of forecasting efficiency [44] is 53 for the women and 24 for the men. The former percentage is well above the forecasting efficiency reached by the best of the mental tests, and the latter is about as good as the average mental test, according to Hull,[45] who places the useful range of efficiency for mental tests between 13 and 29 percent. No doubt this efficiency could be raised by a better approximation of the curves for hue. But this might bring in more harmonics and hence lengthen the equations. More of the factors aside from hue, tint, and chroma could perhaps be brought into consideration and they could be given weight in the formula. But this again would lengthen the equations.

It must be remembered, also, that the predictions dealt with here are based upon averages. Every individual might have his own equation, with its own constants which differ somewhat from the ones presented here. Even when such constants are known for a given individual, one could not predict what his very next affective reaction to a certain color is likely to be, nor what his average reaction might be to a color when it appeared under conditions other than those observed in these experiments. Hence, a practicable, workable equation for prediction would have to include besides the factors that were studied here, such things as form, background, illumination, use of the color, texture, and perhaps many others.

Under the somewhat isolated conditions of these experiments, in which practically all the visual field except for the patch of color was of a uniform gray, many factors besides hue, tint, and chroma were patently at work. The remaining 29 percent of determination for the women and 74 percent for the men are due to some of these additional factors. Let us attempt to guess what some of these other factors were.

[44] As determined by a formula given by C. Hull in his *Aptitude Testing*, p. 268.

[45] *Ibid.*, 275.

Very likely, as in all such mental measurements, unsystematic, chance, errors form a part of the factors unaccounted for. Personal biases; the so-called *s*-factors discussed in connection with Table I also are factors. Reverting to Table I, it will be noticed that the correlations of the pooled judgments with the *g*-factor were .995 and .855, for women and men respectively. The square of these, in other words .99 and .73, may be taken as measures of the extent, in terms of percent, to which the pooled judgments express the *g*-factor. The remaining 1 percent and 27 percent would indicate the extent to which individual idiosincracies entered into the pooled judgments of this study. In the case of the women, then, if the judgments are properly weighted in the pooling, we would expect to find that the factors additional to those of hue, tint, and chroma, are determiners held in common by all five individuals. But in the case of the men, many of the extra determiners are probably in the nature of personal idiosincracies, and perhaps also in the nature of poor affective discrimination.

Among the systematic, universal determining factors, over and above the three attributes of color, one of them is probably the preference for the 'unmixed' colors. This, however, is not nearly as evident in Fig. 2 as in Fig. 1. And following the thorough analysis which has been made, we find three maxima coming exactly at three of these so-called unmixed colors, leaving only yellow not taken care of. The apparent advantages for orange and violet have all but vanished in Fig. 2.

Other systematic deviations may be sought in a study of the discrepancies, especially among the larger ones, in columns z' of Table IV. It will be noticed that wherever z' is relatively large, greater than .50, for example, for one sex, it is likely to be large and of the same sign for the other sex. There are fifteen such cases. There are other cases in which the discrepancies are smaller but of like sign for the two sexes. In fact the coefficient of correlation between the two sets of discrepancies (z') for men and women, is .803. The correlation between the original AVs for the two sexes is only .607. These two coefficients taken together indicate strongly

that the two sexes agree better with one another on the effects of factors other than hue, tint, and chroma, than they do upon the relative effects of *all* common factors *including* those three attributes. Such an agreement surely means additional common factors. But what the factors are, the writer has not been able to ascertain. Many of them are probably social factors—fashions, fads, and other results of social conditioning. After we have ruled out the effects of hue, tint, and chroma, and of any other more deep-seated causes of color preferences and have then found consistent discrepancies, we can hope to identify the consequences of such more superficial factors as social conditioning, or 'mere conventions.'

Color Preferences and Color Vision.—We come now to the theoretical significance of the periodic functional relationships which were assumed to exist between hue and affective value. Taking the curves in Fig. 2 just as they stand, we note that the three maxima are at red, green, and blue. The three minima are at yellow, blue-green, and red-violet. For every hue in the first group, there is a complementary hue in the second group. Furthermore, the three maxima come close to the three primaries of classical color theories, except that the 'blue' of the theories is farther over into the violet. It will be noted that the three maxima differ in height. The general rule is, the shorter the wave length, the higher the maximum. The three minima do not quite follow this same rule, the middle minimum being higher than the other two. Many more points of significance might be pointed out in the total curves.

Much greater significance may be seen if a curve is analyzed graphically into its two major harmonics, as in Fig. 3. Here only the women's curve is drawn, since it is based upon more reliable data and it is probably more representative of the general law for hue. Each harmonic, although represented by a 'sin' and a 'cos' in the original analysis of Table III and in the formulas, is reducible to a single sine curve of different 'phase' simply by summing the two components, as has been done in Fig. 3. We can now note where the

maxima and minima of the two harmonics fall and see whether or not those points have any important meaning.

Now it will be seen that the first harmonic merely has its one maximum at blue and its one minimum at yellow, or at a slightly orangish yellow. This might have been expected from the shape of the original curve. That they come exactly at complementary hues should not be surprising, either. The hues have been evaluated on a scale which puts complementary colors at opposite poles of the color circle. But of

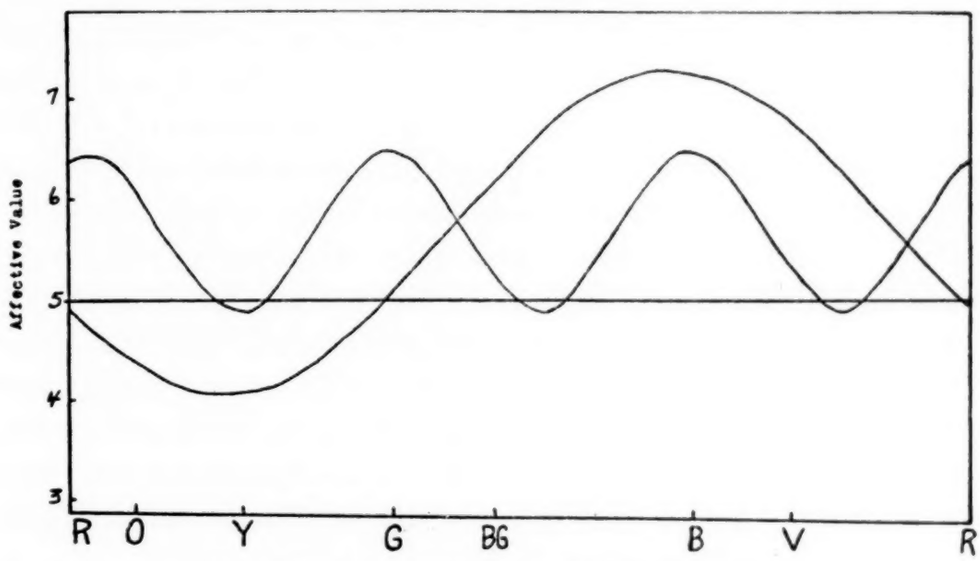


FIG. 3. The first and third harmonics in relation to the various hues.

all hues where the maximum and minimum of the first harmonic may have chosen to fall, why did they select blue and yellow?

The third, and other important harmonic in Fig. 3, has three maxima, at red and green as well as at blue. Why, of all hues, did the third harmonic select these choice positions for its three maxima? I am now going to suggest the hypothesis that *these two harmonics actually represent two different systems of color appreciation.* And the further suggestion comes easily: that *these two systems are bound up with two corresponding systems of color vision.* They are the two evolutionary systems or stages so long championed by Ladd-Franklin. As applied to color preferences, the evolutionary theory would

mean that with the first yellow-blue discrimination the short waves aroused agreeable affective processes and the long waves a disagreeable state of feeling. In so far as the primitive yellow-blue system still operates, the fundamental color discrimination carries with it the corresponding affective discrimination. Occasionally a writer has sensed this distinction between short and long wave stimuli, or between cool and warm colors. Walton [46] has recently found that albino rats show a consistent and general preference for blue and green over yellow and red. I do not maintain that the albino rat is still in the yellow-blue stage of evolution, for Walton's results give very strong evidence that the rat sees red as a color. However, this would not be necessary. Even in human subjects with full color discrimination, the yellow-blue system is still the stronger one for color preference.

As the final red-green distinction came, not all the long wave stimuli remained unpleasant. The two extremes of this range of wave lengths became pleasing along side the blue. Or, to put it in another way, Ladd-Franklin's most recent color scheme includes three primaries, all of which, being primaries, relatively pure, unmixed and well saturated in appearance, are pleasing. Yellow, being a mixture, physiologically, in this new scheme, gains nothing in affective value.

Like much theorizing, perhaps this discussion has led too far beyond its factual basis. But even false theories are useful if they lead to new attacks upon a problem. The relations pointed out here between the agreeableness of colors and theories of color vision can be put to experimental tests in a number of ways. Methods already yielding facts about color vision itself may be duplicated in dealing with color preferences to see whether corresponding results are obtained. For example, the writer has already begun a study of color preferences in relation to color blindness. Such studies of preference might not only reveal something about the factors in the affection for colors, but they might also throw some light upon the nature of color vision itself.

[46] W. E. Walton, color vision and color preference in the albino rat, *J. Comp. Psychol.*, 1933, **15**, 373–394.

Summary and Conclusions

Forty colors, evaluated for hue, tint, and chroma on the Munsell scheme, were judged repeatedly for affective value by five men and five women subjects. With tint and chroma held fairly constant, the functional relationship between affective value (AV) and hue appeared to be of a periodic type. Fourier analysis revealed two important harmonics (first and third) in this relationship, and two others of less important (second and fourth). With hue held constant, linear relationships were found to hold between AV and tint and chroma. Complete empirical formulas were set up, predicting AV from the three attributes of color. The following conclusions have been reached, subject to the usual qualifications, 'under the conditions of this study':

1. Hue determined AV to the extent of about 67 percent for the women and 16 percent for the men. This is based upon a periodic functional relationship involving not more than four harmonics.

2. Tint determined AV about 20 percent for the women, but only 5 percent for the men. In both cases, the lighter the tint the more pleasing the color, all other things being equal.

3. Chroma determined AV only about 5 percent for the women but 13 percent for the men, the more saturated colors being preferred to the less saturated.

4. The three attributes combined account for at least 71 percent of the AV of color for women and 26 percent for the men.

5. There are other common factors, in addition to hue, tint, and chroma, that determine the AV of colors. These can be identified after the effects of hue, tint, and chroma have been eliminated.

6. Some theoretical consequences of the periodic relationship between AV and hue are:

(1) There are two distinct systems of color preference, the one being more primitive and fundamental than the other.

(2) The more fundamental system is a yellow-blue affective discrimination, the short light waves being more agreeable than the long.

(3) The other system is a red-green-blue preference over yellow, blue-green, and red-violet.

(4) These two systems correspond in a striking way to the two evolutionary stages of the Ladd-Franklin color theory.

(Manuscript received March 29, 1933)

THE INTERRELATIONS OF SPEED OF REACTION MEASUREMENTS[1]

BY LYLE H. LANIER

Vanderbilt University

The study of the interrelations of speed in different types of activities has developed during the past decade into an important experimental problem in psychology. The analysis of the nature of behavior organization by the correlation method, associated most prominently with the name of Spearman, has been extended to include many different 'factors,' among the most prominent of which is speed. Phenomenal events occur 'in time,' and rate of response is perhaps the one universal measurable characteristic of reactions. It is only natural that attempts should have been made to determine the extent to which this speed variable represents a 'common' organic factor conditioning diverse types of behavior.

Wissler, in 1901 (**21**), reported the first quantitative correlations on this question, and although his technique was probably inadequate to the task of establishing his general conclusions, he nevertheless developed an important problem and method for psychology. A year later, Aikens, Thorndike and Hubbell (**1**) published results which seemed to support Wissler's general conclusions, indicating that measurements of speed of reaction in several 'perceptive and associative processes' were essentially unrelated. This was

[1] The present article is based upon the results of four experiments carried out by the writer and by certain of his graduate students. Acknowledgments are gratefully made to Mr. John Leedy (9), Miss Martha Lambeth (8), and Mr. A. W. Pierce (11) for the results reported in the second, third and fourth experiments, respectively. These studies dealt with other problems than the analysis of the general nature of 'speed', but only the data in them bearing upon this question are incorporated in the present article.

The writer has in preparation a general article on this subject which contains a rather detailed review of the literature. Consequently, no attempt is made here to do more than refer to certain results of previous investigations.

perhaps the first expression of Thorndike's theory of the 'specificity of mental functions,' a view which has been opposed by Spearman's theory of a general factor. The specific problem of speed received only incidental attention during the period 1902–1924, the burden of interest centering on the question of the 'nature of intelligence,' or of the organization of 'abilities' in general. It is true that Hart and Spearman in 1914 (**6**) reported results designed to demonstrate that 'speed' was not a 'group factor' unrelated to 'intelligence,' but their choice of tests and methods of isolating speed experimentally were open to criticism. This study is important, however, in that it defined the problem of 'speed.' It was inevitable that active experimental interest should develop, both on account of the intrinsic importance of the problem and because of the objectivity of time as a unit of measurement or symbol of reaction.

Omitting the several studies of 'speed' versus 'power' as indicators of 'intelligence,' in relation to group intelligence test work, it appears that Bernstein (**2**) was one of the first investigators to study intensively the question of the interrelations of speed in different activities. His methods and results will be considered somewhat in detail in the article by the writer already referred to, and this review will not be duplicated here. Bernstein concluded against the existence of a 'speed' factor independent of 'intelligence,' holding in substance that an individual who is fast in one operation will not necessarily be fast in another. Experimental studies of the problem since Bernstein's work have been rather numerous and have been increasingly characterized by a more careful control of experimental conditions. These studies have, nevertheless, yielded quite a diversity both of data and of interpretation. Certain investigators have reported high correlations among measurements of speed in different processes, and between speed in simple motor reactions (even reflexes) and 'intelligence' test scores (**10, 12, 14, 18**). At the other extreme are studies which fail not only to find any relationship between speed of reaction and intelligence, but also between speed measurements in different, rather simple motor performances (**4, 15, 16, 19, 20**).

Although the present writer inclines to the view that the general trend of the experimental work supports the second hypothesis, nevertheless, certain results are incompatible with it. McFarland (**10**), for example, in a well-controlled experiment reported unusually high correlations between simple auditory reaction time and rate of response in several 'mental' reactions (*e.g.*, verbal discrimination and association reactions). Similarly Miss Kennedy (**7**) found a correlation of .54 between Army Alpha test score and a composite score based upon seven 'speed' functions, although no correlation resulted when the Army Alpha score in *unlimited* time was used. Sisk (**16**) and Farnsworth, Seashore and Tinker (**4**) failed to find any relationship between simple speed processes and intelligence test score. Furthermore, both of these studies found only a low, unreliable relationship between speed in simple and in more complex motor activities. These studies contradict both the view that 'speed' depends upon a somewhat unitary function closely related to 'intelligence' (Peak and Boring, Rounds, McFarland), and the theory that 'speed' is based upon a 'group factor' essentially unrelated to 'intelligence' (Miss Kennedy). Tinker interprets the experimental results on this problem to indicate definitely that the "average intercorrelation between performance in motor tests of various kinds is approximately zero, and hence motor performance is not appreciably affected by any general factor" (**17**). He notes further that comparatively high correlations may be found between similar kinds of performance, a point which will be emphasized below in connection with results secured in the present experiments. With respect to the much discussed question of 'speed' versus 'power' as an index of intelligence level, Tinker observes that these two types of symbols of efficiency are practically equivalent for the same kind of material.

It should perhaps be noted that such discrepancies in experimental results as those mentioned may be due to several factors, among which the following would seem to be the most important: (1) unreliable data, due either to too few subjects or to an insufficient number of measurements; (2) differences

in the range of processes studies; (3) experimental errors; (4) differences in the homogeneity of the subjects used. The experimental results to be reported below may by no means be free from all of these possible sources of error. The four experiments do represent, however, somewhat diverse approaches to the general problem of the interrelations of 'speed' in different types of activities. In the first experiment a wide range of different types of activities (not merely 'tests' with different names) was included, with a view to determining whether or not the interrelationships found would reveal any consistent trends. Too often in previous work too few activities have been studied, or the activities were obviously too similar to justify conclusions about 'speed' in general. The second experiment was quite similar to the first, except that fewer activities and a larger, more homogeneous group of subjects were used. As in practically all previous studies of speed, the subjects were college students, but these subjects were undergraduate men within a very narrow age range. The third experiment involved the use of twelve-year-old white and negro boys. Five of the speed tests used in the second experiment were used, with certain other tests, in this experiment. The groups of subjects were carefully selected so as to secure representative samplings (as indicated by school grade status) in each race. They would be expected to include a relatively much wider, more heterogeneous range of 'ability' than the college groups usually studied. The fourth experiment was a study of the relationships of speed in activities known rather definitely to be integrated at different neural levels: simple reaction time, the lid reflex, the psychogalvanic reflex and intelligence test performance. This experiment was designed to test the hypothesis that rate of nerve conduction is the essential factor in speed of reaction, and that this factor conditions speed in diverse activities somewhat uniformly. These varied techniques and types of subjects should yield results which would throw considerable light upon the contradictions found in previous experiments.

I. The First Experiment

1. *Subjects and Methods of Procedure*

Thirteen students in a graduate course in experimental psychology were used as subjects for the experiment, which was carried out as a part of the laboratory work of the course. The group consisted of six men and seven women, varying in age from 23 to 46. Inasmuch as the content of the course during the first six weeks of the term (summer of 1927 at George Peabody College) consisted exclusively in the study of speed of reaction measurements, it was determined to outline the laboratory work on a project basis, with the general aim of studying the interrelations of speed measurements in diverse types of activities. All the students were graduate students, had had several courses in psychology, and most of them had done considerable experimental work. Seventeen types of reactions were selected for study and these were divided into four groups. The processes studied are listed in Table 1, being numbered from 1 to 17. The four groups of activities included, respectively, the following factors: (1) 1, 2, 3; (2) 4, 5; (3) 6, 7, 8; (4) 9–17 inclusive. The thirteen students were divided into four corresponding groups, the size of each varying with the time required for the measurements. Two periods were spent in outlining carefully the methods to be used. In addition, the instructor usually gave the experiments to one member of each group and observed this individual act as experimenter for at least one subject. In this way it was hoped that the variations and errors from lack of practice and diversity of experimenters would be largely obviated. Furthermore, inasmuch as the measures were to be used only for the purpose of differentiating the several subjects, it was thought that such results might throw significant light on the general speed interrelationships involved, irrespective of possible experimental errors which might have affected the validity of the absolute time values secured. The methods employed in measuring speed in the different activities will now be briefly described. Inasmuch as several of the reactions studied in this experiment were used also in later experiments under the same conditions, the methods will be described somewhat in detail here.

Simple reaction time.—A Dunlap chronoscope and accessories were used in the measurement of all three kinds of simple reactions (visual, auditory, tactual). The pneumatic reaction key was used by the subject, who was uninstructed as to whether or not he should adopt the 'sensorial' or the 'muscular' set. All of the simple reactions were measured at one sitting, in the following order: 20 auditory, 20 visual, 40 tactual, 20 visual, 20 auditory. A varying interval (1 to 3 seconds) elapsed between the 'ready' signal and the stimulus.

Discrimination reaction time.—Both the 'discrimination' and the 'discrimination and choice' reactions were measured with the Marietta chronoscope. For the first of these two the subject was told that either a green or a white light would be flashed in the stimulus lamp. He was instructed to press his key (a telegraph key) when the green light appeared, but not to react to the white light. One hundred stimuli were given, 50 green and 50 white, each preceded by a warning signal. Thus fifty reactions were measured. ('false' reactions were noted but no time record was kept for them).

The directions were slightly different for the 'discrimination and choice' reactions. The subject reacted with the right hand if the green light appeared and with the left if the white light was presented. Fifty such reactions were measured, equally divided between the two hands.

Association reactions.—Three types of association reactions were measured: (1) free association, in which the subject reacted to the stimulus word with the first word

which occurred to him, (2) partially controlled association, in which the subject responded with a word beginning with the same letter as the stimulus word, (3) fully controlled association, in which the subject gave an antonym of the stimulus word. Twenty reactions of each type were measured. These reactions were timed with a fifth-second stop watch.

Speed of serial movement.—One of the Minnesota mechanical abilities tests was used here. The subject marked vertical lines between the horizontal lines of a slender rectangular figure which extended across the test sheet. Three trials were made. The average time per reaction in milliseconds was computed by dividing the time per trial by the number of reactions made on the 'median' trial.

Speed in cancellation.—This cancellation test was also developed at the University of Minnesota. The test sheet consisted of 1000 capital letters arranged in 26 rows. Two hundred of them were 'A's,' which the subject was instructed to mark out. Two minutes were allowed for the test and the average time per reaction was calculated, as above, by dividing this time by the number of A's crossed out. Errors and omissions were negligible and were disregarded in the scoring.

Substitution.—A digit-symbol test developed at the University of Minnesota was used. Four minutes were allowed and the number of substitutions divided into this time gave the average time per reaction used in the correlations. No subject finished the test.

Simple card sorting.—The subject distributed a deck of playing cards into four piles according to suit. Three trials were made. The 'median' of the three time values was used in computing the average time per reaction.

Complex card sorting.—A deck of flinch cards was sorted in a sorting box of three rows, with five compartments to the row. There were thus fifteen compartments, which corresponded to the numbers on the cards. The compartments were labeled by fastening a card, of the same sort as those to be distributed, to the back of the compartment in plain view of the subject. Three trials were allowed and the median time selected as the subject's speed rating for this process.

Color naming: colored squares.—The Woodworth and Wells sheet containing 100 small colored squares (Bk, Y, G, R, B) was handed to the subject. He was directed to call out the names of the colors as fast as he could. Three trials were made, with a rest period between each. As before, the median time was used to obtain the average time per reaction.

Color naming: printed names of colors.—A sheet containing the printed names of the colors in the Woodworth and Wells test was used. The names were copied in an order the reverse of that of the colored squares. The general procedure otherwise was the same as in the preceding test.

Form naming: geometrical forms.—The subject was instructed to call out as rapidly as possible the names of the geometrical figures on the Woodworth and Wells sheet. The page contains 100 forms. The five forms (circle, square, star, cross, triangle) are printed in random order. The procedure was the same as with the color-naming tests.

Form-naming: printed names of forms.—As with the printed names of the colors, the names of the forms just referred to were copied, in reverse order, and the subject was instructed to call out the names as rapidly as possible. Three trials were made.

2. *Results of the First Experiment*

The average rate of response for each subject in each of the 17 processes was converted into millisecond units. Group averages and measures of variability were computed from these values, and these are shown in table 1. Such figures

TABLE 1

MEDIANS, QUARTILE DEVIATIONS AND 'COEFFICIENTS OF VARIABILITY' ($V = Q/md.$) FOR THE 17 TYPES OF REACTION IN THE FIRST EXPERIMENT
N = 13 GRADUATE STUDENTS (6 MEN AND 7 WOMEN)

Type of Reaction	Median Time (Milliseconds)	Quartile Deviation	'Coefficient of Variability' ($V = Q/md.$)
1. Simple reaction time (tactual)	209	63	.310
2. Simple reaction time (auditory)	230	50	.218
3. Simple reaction time (visual)	226	24	.106
4. Discrimination reaction time	569	65	.114
5. Discrimination and choice time	525	54	.103
6. Free association reaction time	2202	394	.179
7. Partially controlled association time	2555	343	.134
8. Fully controlled association time	2056	397	.192
9. Minnesota 'speed of movement'	175	16	.091
10. Speed in cancellation	790	92	.116
11. Speed in substitution	1757	223	.127
12. Simple card sorting	1040	203	.195
13. Complex card sorting	1835	250	.136
14. Color naming: colored squares	578	102	.177
15. Color naming: printed names	308	30	.098
16. Form naming: printed forms	709	75	.105
17. Form naming: printed names	368	37	.101

based upon only 13 cases are of doubtful value, but they are presented both as a general check on the data from which the correlations were calculated and for comparison with similar figures in the other experiments. The median and quartile deviation, in millisecond units, are shown for each type of reaction. From these two values a crude index of relative variability was computed by dividing the quartile deviation by the median. Similar values are shown for the data of the later experiments, with which these figures will be compared. Except for the long association times and the high degree of variability for the simple reaction measurements, the figures show fairly good agreement with previous results. The long association times are due undoubtedly to the inadequate method of timing the reactions with a stop watch. These

TABLE 2

INTERCORRELATIONS OF SPEED OF REACTION MEASUREMENTS IN THE FIRST EXPERIMENT. ALL COEFFICIENTS OF LESS THAN .55 ARE STATISTICALLY UNRELIABLE (LESS THAN FOUR TIMES THE PROBABLE ERROR)

Factors Correlated	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Simple reaction time (tactual)	.82	.88	.36	.36	.22	−.09	−.12	−.64	−.44	−.48	.03	−.60	−.42	−.04	−.41	−.22
2. Simple reaction time (auditory)		.90	.52	.50	.36	−.14	−.02	−.18	−.32	−.46	.20	−.64	−.45	−.09	−.14	−.38
3. Simple reaction time (visual)			.46	.42	.33	−.38	−.09	−.30	−.64	−.37	.22	−.55	−.35	−.03	−.26	−.42
4. Discrimination reaction time				1.00	.38	−.02	.00	−.44	−.56	−.32	−.20	−.44	−.38	−.44	−.01	−.34
5. Discrimination and choice time					.16	−.19	−.18	−.44	−.19	−.34	−.38	−.61	−.49	−.16	−.07	−.30
6. Free Association reaction time						.78	.79	−.08	.33	.06	.30	.02	−.06	−.20	−.19	−.56
7. Partially controlled association time							.91	.01	.11	.35	.08	.14	−.38	−.50	−.33	−.54
8. Fully controlled association time								.14	.18	.48	.05	.32	−.19	−.17	−.03	−.33
9. Minnesota 'speed of movement'									.00	−.20	.48	−.10	.00	.14	.14	.06
10. Speed in cancellation										.60	.38	.59	.20	−.02	.08	−.24
11. Speed in substitution											.26	.66	.42	.20	.38	.06
12. Simple card sorting												.12	.16	−.15	.14	−.52
13. Complex card sorting													.67	.17	.36	.09
14. Color naming: colored squares														.53	.69	.44
15. Color naming: printed names															.57	.81
16. Form naming: printed forms																.39
17. Form naming: printed names																

measurements would have been discarded entirely but for the fact that the internal consistency manifested by their interrelations indicated that the errors in timing were probably of the chance variety. Individual differences in association time seem to have been adequately manifested under the conditions of the experiment.

The subjects' ranks in each function were determined and from these values the correlation of each factor with every other was computed by the rank-difference method. The resulting 136 coefficients are shown in Table 2. It will be noted from the names of the variables correlated that certain groupings have been attempted in order to bring together in the table coefficients representing correlations of apparently related activities. In general, the first 8 activities are 'discrete' reactions to isolated stimuli, whereas the last 9 are 'serial' motor reactions. Within the 'discrete' group will be found three sub-groups: (1) 'simple' reactions; (2) discrimination reactions; (3) association reactions. Among the serial activities the four 'naming' reactions constitute a fairly distinct group and these are placed at the end of the table for the sake of easy inspection. The two card-sorting reactions are placed together. Of the remaining three reactions, cancellation and substitution have certain features in common (this judgment is supported by the correlation of the .60 found for them), and they are placed in contiguous positions in the correlation table. The serial movement test involves no discrimination whatever and would appear to stand somewhat alone in the series of reactions used.

TABLE 3

AVERAGES OF CERTAIN CORRELATIONS REPORTED IN TABLE 2

Average intercorrelation of simple reaction factors	.866
Correlation of the two discrimination reactions	1.00
Average intercorrelation of association reactions	.826
Average correlation of simple and discrimination reactions	.44
Average correlation of simple and association reactions	−.008
Average correlation of discrimination and association reactions	.025
Average intercorrelation of all 'discrete' reaction times	.32
Average intercorrelation of 'naming' reactions	.57
Average intercorrelation of serial manual reactions	.29
Average intercorrelation of all 'serial' reaction times	.24
Average correlation of discrete with serial action times	−.23

A study of the coefficients in Table 2 will indicate that the grouping on the basis of apparent similarity in 'set' and effector mechanism was not without basis as regards real relationships. To facilitate the analysis of Table 2, Table 3 has been prepared, in which certain averages are shown. Thus one notes that the average intercorrelation existing among the three simple reaction measurements was .866. The two discrimination reactions correlated perfectly (*r* equals 1.00). The three types of association reactions also showed a high degree of relationship, the average intercorrelation being .826. The relationships among the reactions within each sub-group in the 'discrete' category were, then, exceptionally high. Average correlations between factors in these three sub-groups were much lower. Simple reaction time gave an average correlation of .44 with discrimination reaction time. With association time, the average correlation was −.008, although this average obscures the fact that the correlation of simple reaction with free association time was higher (.22, .36, .33) than with the controlled association times, where all of the coefficients were negative. Likewise discrimination reaction time showed no correlation with association time, the average *r* being .025. In general, it would appear from these results that where the 'set' and effector mechanisms are similar, correlations of reaction time measures tend to be high, but as the reactions correlated involve greater and greater postural and effector diversity the correlations of speed measurements become less and less.

The analysis of the correlations among the serial action factors supports this hypothesis, although the reactions in this group are less susceptible of grouping than were the discrete reactions. The four 'naming' reactions yielded an average intercorrelation of .57, the coefficients ranging from .39 to .81. The latter value was secured between reading the printed names of geometrical forms and of colors, two reactions which are obviously similar. The reading of the forms and of the colors themselves present situations which are rather more 'difficult' and unusual. These two factors correlate rather highly, .69, although the correlations of each with the two other naming reactions are appreciably lower.

Factors 10–13 inclusive are serial manual reactions of varying degrees of complexity. Being heterogeneous with respect to difficulty, there would perhaps not be expected such a high degree of relationship as was found among the similar groups of reactions already studied. The average correlation found for this group of factors was .29. However, if one omit factor 9, which involved merely speed without any discrimination, the average correlation among the remaining four reactions is .44. Cancellation, substitution and the two card-sorting tests are serial reactions apparently involving discrimination and selectivity of response. The moderate degree of correlation found suggests that these common factors are not merely apparent but actually exist. In fact, if one carry the process of elimination further and omit simple card-sorting from the average, this average is raised to .62. The simple card-sorting activity requires very little discrimination; the belief that it more nearly involves simple manual speed of movement is supported by the correlation of .48 with 'speed of serial movement' (factor 9). These results support the conclusion reached regarding the interrelations found among the reactions in the 'discrete' class. Similarity in 'posture,' in the type of movement required, in the general perceptual-motor situation presented, result in an order of positive correlation among speed measurements commensurate with the degree of such similarity.

The average intercorrelations of all of the discrete reactions, of all of the serial activities, and the average correlation of discrete with serial reactions are shown also in table 3. These figures would seem to contradict the theory of a general 'speed' factor conditioning rate of response in diverse types of reactions. Such averages really have little analytical meaning, but they serve to emphasize the important fact that before one can speak of the relation of 'speed' to 'mental ability' or to any other process, the problem of what 'speed' means must be settled. These results manifest an unusual degree of internal consistency and perhaps on this account one can place more reliance on them than the small number of subjects might otherwise merit. The figures cited support

those of Sisk and of Farnsworth, Seashore and Tinker, in so far as those studies are comparable with the present one.

II. The Second Experiment

1. *Subjects and Methods of Procedure*

The data to be presented here were collected as a part of a research on the relationship between 'speed of reaction' and basal metabolism, carried out by Mr. John L. Leedy under the writer's supervision (9).[2] Inasmuch as it has already been suggested that the term 'speed of reaction' has no general, unitary meaning, it might appear that work on such a problem would be futile. The use of this generic term did not, however, imply any definite theory as to the behavior organization concerned. Speed in some ten rather simple (but not uniformly so) discrete and serial motor reactions was measured, the aim being primarily to use these measurements as the basis for selecting a 'fast' and a 'slow' group for the basal metabolism comparisons. These ten types of reactions, listed in Table 4, also provided considerable material for correlation analysis. It is this aspect of the results that will be considered here. Eight of the ten types of reactions were studied also in the first experiment, and no description of the detailed methods need be given for these processes (the three simple reactions, discrimination reaction, Minnesota speed of movement, cancellation, color naming, naming names of colors). The only variations in procedure were the use of the Bergstrom chronoscope to measure the times of all discrete reactions, and the measurement of 50 simple reactions of each type instead of 40.

Rate of stylus tapping and one of the Minnesota paper tapping tests were the two speed tests not used in the first experiment. In the first of these the subject tapped with the stylus for 15 seconds, the number of taps being recorded by an electric counter. Three trials were made. The Minnesota tapping test blank contains six blocks of 5-millimeter squares. The subject was instructed to place one dot in each square as rapidly as possible. Five trials of 15 seconds each were made. In addition, the Army Alpha Test, Form 5, was given to all of the subjects.

Thirty-four undergraduate men were used as subjects in the experiment. The group was, accordingly, homogeneous as to sex and age, in contrast with the subjects in the first experiment.

2. *Results of the Second Experiment*

Measures of central tendency, variability and 'relative variability' for all types of reactions used in this experiment are shown in table 4. A comparison of the averages with those of the same processes in the first experiment indicates that the present subjects were considerably faster in all of the discrete reactions, but were slower in all of the serial activities except color naming. In regard to variability, the coefficients of relative variation show that generally the 13 subjects in

[2] The specific results of the basal metabolism study will be reported in another article. They are not directly related to the present problem, except perhaps as regards the question of the organic basis of 'speed' in different activities.

the first experiment were more variable than the present subjects. This is especially true in simple reaction time. When one considers the constitution of each group, this variability

TABLE 4

MEDIANS, QUARTILE DEVIATIONS AND 'COEFFICIENTS OF VARIABILITY' ($V = Q$/median) FOR THE MEASUREMENTS INCLUDED IN THE SECOND EXPERIMENT

	Median Time (Milliseconds)	Quartile Deviation	'Coefficient of Variability' ($V = Q$/md.)
1. Simple reaction time (auditory)	146	13.0	.089
2. Simple reaction time (visual)	180	20.5	.114
3. Simple reaction time (tactual)	180	16.0	.089
4. Discrimination reaction (visual)	375	35.5	.095
5. Stylus tapping	135	11.0	.081
6. Minnesota speed of movement	190	12.5	.066
7. Minnesota tapping test	326	22.0	.067
8. Naming names of colors	360	43.8	.122
9. Color naming	522	50.5	.097
10. Cancellation	895	72.5	.081
11. Army Alpha	148 (points)	12.2	.082

seems natural enough. The 13 graduate students included both sexes and the range in age was from about 23 to 46. The subjects in the second experiment were all male undergraduates within quite a narrow age range. This difference in homogeneity would seem to be an important factor in accounting for certain discrepancies in the correlation results of the two experiments, which will be considered below.

Intercorrelations of the ten 'speed' factors and of Army Alpha scores are shown in table 5. A study of the coefficients shown there reveals agreement in certain respects with comparable results of the first experiment. The highest correlations are seen to exist between the three types of simple reaction measurements (.78, .64 and .76), the average correlation being .72, as compared with .866 in the preceding experiment. The three simple reaction times gave an average correlation of .45 with discrimination time, as compared with .44 in the first experiment. The average inter-correlation among the serial reactions was .38, as compared with .24 in the first experiment (where, incidentally speed in nine instead of six serial activities was measured). The fact that the speed tests in this experiment were 'simpler,' and perhaps

TABLE 5

INTERCORRELATIONS OF SPEED OF REACTION MEASUREMENTS FOR 34 COLLEGE MEN

The last column shows correlations of the speed measurements with Army Alpha scores, based upon 30 cases

Factors Correlated	1	2	3	4	5	6	7	8	9	10	11
1. Auditory reaction time		.78 ±.041	.64 ±.046	.42 ±.096	.44 ±.091	.23 ±.110	.35 ±.104	.23 ±.110	.25 ±.110	.44 ±.091	.32 ±.112
2. Visual reaction time			.76 ±.050	.54 ±.082	.31 ±.104	.23 ±.110	.50 ±.082	.37 ±.096	.35 ±.104	.46 ±.082	.46 ±.092
3. Tactual reaction time				.40 ±.096	.17 ±.110	.27 ±.104	.31 ±.104	.19 ±.110	.08 ±.113	.42 ±.096	.40 ±.103
4. Visual discrimination time					.25 ±.110	.33 ±.104	.33 ±.104	.10 ±.113	.19 ±.110	.10 ±.113	.22 ±.128
5. Stylus tapping						.54 ±.082	.24 ±.104	.40 ±.096	.40 ±.096	.31 ±.104	−.06 ±.122
6. Minnesota speed of movement							.44 ±.091	.44 ±.091	.40 ±.096	.50 ±.082	.21 ±.118
7. Minnesota tapping test								.15 ±.110	.19 ±.110	.37 ±.096	.23 ±.118
8. Naming names of colors									.37 ± .096	.44 ±.091	.46 ±.092
9. Color naming										.48 ±.082	.10 ±.121
10. Cancellation											.19 ±.119
11. Army Alpha											
Correlation of Army Alpha with composite speed rankings											.29 ±.112

somewhat more homogeneous than those in the preceding one, might appear to explain the somewhat higher correlations secured here. However, of the reactions omitted from the present experiment, substitution and complex card-sorting correlated rather highly (.66) in the first experiment, and both of these factors gave high correlations with cancellation (.60 and .59). These facts contradict the hypothesis that the omission of these tests in the second experiment resulted in an increase in the average inter-correlation of the serial action measurements. The discrepancy between these average correlations is not great, however, especially in view of the small number of cases used in both instances. The average correlation of the four 'discrete' with the six 'serial' reactions was .29, a figure comparable in size, although reversed as to trend, with the coefficient of −.23 found in the first experiment. Here again the small numbers of cases and their homogeneity may account for the discrepancy. In general, there is rather good agreement between the results of the two experiments with respect to the types of relationships displayed. The correlations between speed in similar activities tend to be high, while there is a progressive decrease in correlation with increase in postural and effector diversity.

The most surprising result in this experiment is the higher correlation between simple reaction time and Army Alpha score than between the latter and the other 'speed' factors. One would expect a higher correlation between Army Alpha score and the more complex (?) serial action tests than with simple reaction time (*c.f.* Sisk, **16**; Hansen, **5**; Farnsworth, Seashore and Tinker, **4**). While it may be true that a test such as the Army Alpha is really a serial rate test, for a highly selected group of college students, nevertheless the fact remains that in the present experiment a higher correlation was found between this test and simple reaction time than with most of the serial action tests. It is the writer's opinion that such an apparent contradiction is due to the homogeneity of the subjects and to the rather small number of cases. Where a comparatively small, highly 'selected' group is used, accidental errors cause a disproportionate change in the

rankings, with a consequent somewhat radical alteration of the 'correlation' displayed. In Sisk's investigation, where 100 somewhat less highly selected subjects were used, no correlation was found between Army Alpha and simple reaction time. Furthermore, in the fourth experiment of the present series the correlation between Army Alpha and simple auditory reaction time was found to be zero (r equals .00). The subjects were taken from a section of a general psychology class, as in the present experiment. The contradictory coefficients found here may perhaps be interpreted as chance values produced by the effects of accidental 'errors' in sampling or in measurement.

III. The Third Experiment

1. *Subjects and Methods of Procedure*

Except for Bernstein's study practically all of the studies of 'speed' have been made with college students. It has just been suggested that it is difficult to secure a random sampling with such subjects—even if one use large numbers of cases. The present experiment involved the measurement of speed in several types of reactions, using as subjects white and negro twelve-year-old boys. With such subjects it appeared that additional light might be thrown upon the problem of the nature of 'speed.' The analysis of the race differences in performance have been reported elsewhere (8), and this aspect of the results will not be considered here.

Thirty white and thirty negro twelve-year-old boys were selected from certain Nashville, Tennessee, public schools, in such manner as to secure samplings whose grade distributions corresponded almost exactly to those for the city as a whole. For example, the percentages of twelve-year-old white boys in all grades for the entire city were determined. The sampling used in the present study was selected so as to include almost these same percentages of subjects in each grade. Although 30 subjects may appear to be quite a small number, when it is remembered that all of these were boys within an age range of one year, the 'size' of the sampling assumes larger proportions than this number would imply for a group involving both sexes and several ages.

The processes studied are listed in table 6. Five of the 'speed' tests were used also in the second experiment (factors 4, 5, 6, 7 and 8), under the same conditions as prevailed there. The new variables introduced were: (1) the Stanford-Binet Scale; (2) Rational Learning (both the total time required to learn the problem and the average rate of response on each item) and (3) Free Association (the test in the Stanford-Binet Scale requiring the subject to name all the words he can in 3 minutes). The Stanford-Binet Scale needs no description, and the Rational Learning test devised by Peterson has been described in several previous articles (13, pp. 26–29). The subject is required in this test to associate the numbers 1, 2, 3, 4, 5, 6, 7, with the letters A, B, C, D, E, F, G, the numbers having been assigned to the letters in a random order. The experimenter calls out the letter and the subject guesses numbers until he calls the correct one. This procedure is repeated until all of the numbers are correctly associated with the respective letters. In computing 'rate of response' on this test the

TABLE 6

MEDIANS, QUARTILE DEVIATIONS AND 'COEFFICIENTS OF VARIABILITY' ($V = Q/\text{md.}$) FOR THE FACTORS INCLUDED IN THE THIRD EXPERIMENT

The subjects were 30 white and 30 negro twelve-year-old boys. The medians for all 'speed' factors are in terms of average minutes per response.

		Median	Quartile Deviation	'Coefficient of Variability'
1. Stanford-Binet mental age	White	140 mos.	12.96	.093
	Negro	117.5 mos.	13.45	.114
2. Rational learning: total time	White	11.8 min.	5.86	.497
	Negro	15.0 min.	7.62	.508
3. Rational learning: rate of response	White	3.15 min.	.475	.151
	Negro	3.85 min.	.512	.133
4. Minnesota tapping: speed and accuracy	White	.420 min.	.037	.088
	Negro	.480 min.	.065	.135
5. Naming printed names of colors	White	.504 min.	.057	.113
	Negro	.500 min.	.176	.352
6. Minnesota speed of movement	White	.234 min.	.020	.086
	Negro	.235 min.	.035	.149
7. Color naming: colored squares	White	.800 min.	.069	.086
	Negro	.875 min.	.104	.119
8. Cancellation of A's	White	1.40 min.	.215	.153
	Negro	1.63 min.	.260	.159
9. Free Association (no. of words in 3 minutes)	White	2.72 min.	.500	.183
	Negro	3.31 min.	.600	.181

total time for the test is divided by the total number of reactions (correct and incorrect). This total time naturally includes the experimenter's time in presenting the letters, but this is quite uniform and presumably would not affect the relative standings of the subjects.

2. *Results of the Third Experiment*

The averages and variability measures for each of the nine factors included in the experiment are shown in table 6. All of the medians, except those for the first two factors, are expressed in terms of seconds per response. It may be noted, generally, that little or no race 'speed' difference was found for simple motor activities, while the size of the difference increased rather regularly (in favor of the whites) as the complexity of the activities increased. The measures of relative variability in the last column of table 6 show that the negroes generally were slightly more variable than the whites, al-

though the only marked discrepancy between the races in this respect occurred with factor 5 (reading the printed names of colors). The white children used here showed somewhat greater variability (and hence 'heterogeneity') than did the college students on the five tests common to the second and third experiments. The difference was slight, however, since on two of the tests the relative variability was less for the children than for the adults.

The intercorrelations among these 9 factors are shown in table 7. Factors 4–9 inclusive may be considered 'speed'

TABLE 7

INTERCORRELATIONS OF THE NINE FACTORS FOR 30 WHITE AND 30 NEGRO TWELVE-YEAR-OLD BOYS

The upper coefficient in each case is that of the whites; the lower that of the negroes. Probable errors are shown in the lower half of the table. For correlations with factor 5, only 28 negroes were used.

Factors Intercorrelated	1	2	3	4	5	6	7	8	9
1. Stanford-Binet mental age		.51	.15	.05	.27	.27	.27	.70	.30
		.44	.18	.72	.58	.60	.50	.74	.60
2. Rational learning: total time	.09		.11	−.05	.08	−.09	.10	.64	.00
	.10		.50	.17	.27	.18	.21	.23	.28
3. Rational learning: time per reaction	.12	.12		.08	.15	−.06	.24	.06	.19
	.12	.09		−.18	.19	.10	.23	.03	.09
4. Minnesota tapping: speed and accuracy	.12	.12	.12		.27	.41	.05	.12	.27
	.05	.11	.11		.56	.74	.52	.72	.49
5. Color naming: printed names	.11	.12	.12	.11		.31	.41	.22	.34
	.08	.11	.12	.08		.47	.64	.60	.45
6. Minnesota speed of movement	.11	.12	.12	.10	.11		−.03	.05	.18
	.07	.11	.12	.06	.09		.46	.46	.59
7. Color naming: colored squares	.11	.12	.11	.12	.10	.12		.42	−.09
	.09	.11	.11	.09	.07	.09		.73	.45
8. Cancellation of A's	.06	.07	.12	.12	.11	.12	.10		.06
	.05	.11	.11	.05	.08	.09	.05		.59
9. Free Association: no. of words in 3 minutes	.11	.12	.12	.11	.12	.12	.12	.12	
	.08	.11	.12	.09	.10	.08	.10	.08	

variables in the narrower sense of the term. The average intercorrelation among the six 'speed' variables was .27 ± .11 for the whites, and .57 ± .08 for the negroes. The value for the whites agrees fairly well with similar correlations among serial action tests in the first two experiments (.24 and .38), and also with the average intercorrelation of .23 in Sisk's study. The markedly higher average in the case of the negroes is difficult to interpret, especially since they are little

more heterogeneous than the whites. Undoubtedly, some sort of 'common factor' conditions these different speed functions to a much greater extent than is true for the whites. The fact that the average correlation of these six speed tests with Binet mental age was .62 for the negroes would seem to suggest that the latter factor is also an important condition of performance on the speed tests in this group. The implications of these results have been discussed by Lambeth and Lanier, and further comment on them will be made in the next paragraph.

The correlations of Stanford-Binet mental age with these six speed tests were higher (averages of .31 and .62 for the two races) than the intercorrelations among the speed tests themselves. In the case of the whites this higher average is perhaps largely due to the unusually high correlation of mental age with cancellation (r equals .70); the omission of this coefficient from the average reduces the latter to .23. In the case of the negroes, however, the correlations are rather uniformly high, a fact which presents an unusually interesting problem. Lambeth and Lanier point out that these results, taken with others from previous studies, seem to argue for the relatively greater operation of a 'general factor' among the conditions which determine negro performance (**8**). The slightly greater heterogeneity of the negroes, already noted, would seem to be hardly great enough to account for the results. It is conceivable that the presumably more differentiated environment of the white might produce a greater specialization of 'abilities' which would result in a lower average intercorrelation among test scores for that race. Or the negro might be basically of a simpler, more homogeneous, more 'primitive' type of organization. One cannot conclude from these data which hypothesis, if either, is correct, and speculation is useless. With respect to the present problem of the nature of 'speed,' such results are of interest in indicating that the nature of the sampling selected is a most important factor in determining the interrelationships between speed measurements. Such a difference as that found here is sufficiently marked to indicate that the effects upon

correlation results of possible differences in the homogeneity of the sampling and in 'behavior organization' (whether due to innate or environmental factors) should be carefully considered by investigators in this field.

IV. The Fourth Experiment

This experiment differs rather markedly in technique from the other three. The work of Rounds (**14**) and of Travis and Hunter (**18**) produced considerable speculation as to the cause of the remarkably high correlations between reflex speed and speed or efficiency in thinking processes. These investigators, and many commentators, suggest that rate of nerve conduction might constitute the common condition of efficiency in such apparently diverse activities. Although the high correlations of reflex speed and 'intelligence' have not been confirmed by Travis and Young (**19**), nor by Whitehorn, Lundholm and Gardner (**20**), nevertheless the belief in such a basis of a general speed function persists in psychological circles. If such a 'conductivity' factor were really general, one would expect as high (perhaps even higher) correlations between speed in two reflexes as between reflex speed and rate of complex thinking reactions. The present experiment was devised to test this hypothesis and perhaps in that way to throw additional light on the nature of 'speed' of reaction. Lid reflex latency, psychogalvanic reflex speed, simple reaction time and the Army Alpha test score were the factors used in the experiment. Both the lid reflex and the psychogalvanic reflex can be elicited by a loud noise. It seemed highly desirable to secure simultaneous time measurements for the two reactions, inasmuch as gross postural factors would be common to any pair of reactions mediated by the two systems, and their interrelations as regards time might therefore be more accurately determined.

1. *Methods and Subjects in the Fourth Experiment*

The noise stimulus was produced by a large rat trap, similar to that used by Dodge, which was adapted and wired so as to close two separate electrical circuits when the base was struck by the wire spring. One of these circuits included the electro-magnet which released the pendulum of the Bergstrom chronoscope (used to measure the lid

reflex latency). A special switch was constructed and mounted on a spectacles frame for connection with the subject's eye-lid. The switch consisted of the pivotal part of a 'heart lever,' through which ran a small copper wire. One end of the wire was fastened to the eye-lid of the subject by a bit of adhesive tape, while the other end made a contact with a small copper bar which was insulated from the rest of the switch. When the subject blinked, an electric circuit which included a polarized relay was broken. The armature then swung to the opposite pole of the relay, closing a circuit through the pointer magnet of the chronoscope. The pointer, which had been swinging with the pendulum since its release by the stimulus contact, was stopped and the lid reflex time read directly from the millisecond scale over which the pointer had passed until being stopped by its electro-magnet.

The second electric circuit closed by the spring of the trap included an electric marker which recorded the point of stimulation on the smoked paper of an extension kymograph. This kymograph set-up was used to measure the latency of the P. G. R. (psychogalvanic reflex). One of the two experimenters watched the needle of the Hathaway galvanoscope and, when the deflection began, pressed a key which was wired in series with a second electric marker. A Jacquet chronometer made the time record on the smoked paper, in fifths of a second. This is rather a crude method of measuring the time of this response, but since its latent period is quite long (the median time by this method being 2.19 seconds) it was considered adequate for the purpose of revealing individual differences. A repetition of the experiment, using a photographic recording device, is being planned.

Fifty stimulations were given to each subject. The number of reactions per subject varied considerably, since either one or the other reflex, or both, might fail to occur. Table 8 shows that the median number of lid reflexes was 28, while the median number of psychogalvanic reflexes was 31.

TABLE 8

MEDIANS, QUARTILE DEVIATIONS AND COEFFICIENTS OF VARIATION FOR THE SEVERAL FACTORS INCLUDED IN THE FOURTH EXPERIMENT

N = 25 undergraduate men in an introductory psychology class.

	Median	Quartile Deviation	'Coefficient of Variability'
1. Simple reaction time (auditory): millisecond units	201	25.34	.127
2. Lid reflex time: millisecond units	47	7.00	.149
3. Psychogalvanic reflex time: seconds	2.19	.17	.078
4. P. G. R. deflection: milliampere units	.12	.05	.417
5. Number of P. G. R.'s in reaction time experiment	43.25	12.10	.279
6. Number of P. G. R.'s in lid reflex experiment	31.00	11.69	.354
7. Number of lid reflexes	28.50	11.32	.397
8. Army Alpha test	158	14.09	.089

Fifty simple auditory reaction time measurements were also made for each subject, in order to compare speed in the cerebro-spinal and autonomic reflexes with speed in a simple 'voluntary' reaction, presumably integrated primarily at the cortical level. As a check on the psychogalvanic 'reactivity' of the subjects, the Hathaway galvanoscope was connected to the left hand and the number of psychogalvanic reactions occurring as the subject reacted in the 'simple reaction' experiment was noted. The Army Alpha test, Form 5, was given to all subjects.

In addition to the measurements of speed in the three types of reaction, certain other aspects of the reactions were recorded. These include: (1) magnitude of deflection of the galvanoscope pointer, (2) the number of P. G. R.'s in fifty trials in response to the noise stimulus, (3) the number of lid reflexes and (4) the number of psychogalvanic reactions occurring in the reaction time experiment. While these factors may have little or no bearing on the specific question of speed, it was thought that they might throw some light upon the general 'reactivity' of the subjects.

Twenty-five undergraduate men from an introductory psychology class were used as subjects in the experiment. They were closely comparable with the subjects used in the second experiment, although the two experiments were carried out in different years.

2. *Results of the Fourth Experiment*

The averages and measures of variability (absolute and relative) of the eight variables included in the experiment are shown in table 8. With respect to the averages, no very marked discrepancies were found between these and previous results (no comparisons can be made for factors 4, 5, 6, and 7). These subjects are somewhat slower in auditory reaction time than the similar sampling used in the second experiment. The median lid-reflex time of 47 milliseconds agrees well with the results of Dodge (3), despite differences in technique and in the larger number of subjects used here. Dodge found an average lid reflex of about 40 milliseconds, using a photographic method of recording time and reaction. The median time of the P. G. R. likewise agrees fairly well with times reported by other investigators. The Army Alpha average of 158 indicates that this group is rather a select one, as far as performance on this test goes.

Due to the absence of comparative data, not much can be said about factors 4, 5, 6, and 7. The 'magnitude of deflection' value probably has little absolute meaning, due to the many conditions which may affect it. It is interesting to observe that more psychogalvanic reactions occurred in connection with making the 'simple reaction' movement than in response to the loud noise.

The measures of relative variability are interesting in showing rather low variability for the four principal variables in the experiment, while the variability is quite marked for the four incidental factors. The variability coefficients for reaction time and Army Alpha agree well with analogous values in table 4.

The interrelationships among the eight variables are represented by the correlation coefficients shown in table 9.

TABLE 9

INTERCORRELATIONS OF THE MEASUREMENTS MADE IN THE FOURTH EXPERIMENT
N = 25 undergraduate men

	2	3	4	5	6	7	8
1. Simple reaction time (auditory)	.02 ±.135	.19 ±.128	.13 ±.132	.31 ±.122	.02 ±.135	.10 ±.133	.00 ±.135
2. Lid reflex time		−.25 ±.126	.11 ±.133	−.25 ±.126	.04 ±.135	.05 ±.135	−.26 ±.126
3. P. G. R. time			.01 ±.135	.29 ±.123	.36 ±.117	.19 ±.128	.04 ±.135
4. P. G. R. deflection				.11 ±.133	.47 ±.105	.02 ±.135	−.23 ±.126
5. Number of P. G. R., I					.39 ±.114	.00 ±.135	−.17 ±.131
6. Number of P. G. R., II						.13 ±.132	.04 ±.135
7. Number of lid reflexes							.00 ±.135
8. Army Alpha							

The low and unreliable negative correlation of −.25 between time measurements for the two reflexes would seem to indicate that the conditions determining latency in these two reflex arcs tend to vary practically independently. In order to check this relationship further, a coefficient of correlation was computed in which the paired individual latencies (instead of median times for each subject) constituted the items correlated. Thus, instead of 25 pairs of median time values, 415 pairs of individual time measurements, simultaneously secured, were correlated. The resulting coefficient was .06 ± .04, which is probable a truer indication of the relationship than the figure −.25.

The correlations between simple reaction time and speed in the two reflexes were both low and unreliable. On the basis of these results, there would seem to be little or no relationship between speed in the simple reaction, integrated presumably at the cortical level, and speed in either the 'cerebro-spinal' or the 'autonomic' reflex studied here.

The correlations of Army Alpha with both simple reaction time and P. G. R. time are practically zero. A low, negative and statistically unreliable correlation was found between Army Alpha and lid reflex speed (r equals $-.26 \pm .126$). The lack of correlation between speed in these simple processes and intelligence test score is in contrast with that found by Travis and Hunter (**18**), but in agreement with the later work both of Travis and Young (**19**), and of Whitehorn, Lundholm and Gardner (**20**).

The lack of correlation between simple reaction speed and Army Alpha score contradicts the result found in the second experiment (see table 4). Perhaps this fact may throw light on other contradictory results of the same sort in previous studies. Most of the correlations purporting to show any appreciable relationship between simple reaction time and intelligence test score have been based upon comparatively few cases (Peak and Boring, 5; McFarland, 34; Kennedy, 20; Leedy, 35). In such homogeneous, small groups accidental rankings of the subjects might occur which would yield an apparent correlation where no fundamental relationship existed. Sisk's correlations of practically zero between Army Alpha score and simple reaction speed, based upon 100 cases, probably represent more nearly the true status of affairs. The result in the present experiment is in agreement with his figures, and with those of Farnsworth, Seashore and Tinker.

The remaining correlation coefficients in table 9 have no direct bearing upon the problem of speed and they will be given only a slight consideration here. Only the correlation of the magnitude of deflection in the P. G. R. with the number of psychogalvanic reactions in response to loud noise is statistically reliable. This coefficient of $.47 \pm .105$ indicates that the 'intensity' of the P. G. R. tends to be positively correlated with the frequency of its occurrence under similar conditions. This suggests a problem with promising possibilities for further research. The correlation between factors 6 and 7 indicates some relationship between the frequency of P. G. R. under the two conditions obtaining in the experiment (simple reaction situation and the loud noise). More exten-

sive work of this nature might throw light on the range and nature of psychogalvanic 'reactivity.'

Summary and Conclusions

It is rather difficult to draw unequivocal conclusions from the results of four experiments which differ so much either in methods or in subjects. Advocates of all three of the general hypotheses concerning the nature of 'speed' might by judicious selection find apparent support for their respective views in the correlations reported. Therefore, the summary and interpretation of the results to be made below is submitted with full appreciation of the fact that the conclusions reached cannot be regarded as final. Much more careful and intensive laboratory investigation must be carried out before the precise relationships of the conditions determining rate of response in diverse action systems can be known. Nevertheless, certain rather consistent trends are noted in the correlations reported, which, when considered with similar results of previous investigators, would seem to merit the tentative generalizations made below. With these qualifications, the following summary and interpretation of the principal facts may be made:

1. Different types of simple reaction time measurements tend to yield high correlations. The average intercorrelation of simple visual, auditory and tactual reaction times in the first experiment was .866, in the second experiment .72.

2. Two types of 'discrimination' reaction in the first experiment correlated perfectly.

3. Three types of association reactions ('free,' 'partially controlled,' and 'fully controlled') in the first experiment gave high intercorrelations (average r equals .826).

4. Simple reaction time measurements yielded an average correlation with discrimination time of .44 in the first experiment, of .45 in the second.

5. Neither simple reaction nor discrimination reaction time showed any relationship to association time (the average correlations were $-.008$ and .025).

6. Eight 'discrete' reactions (simple, discrimination, and association) in the first experiment yielded an average cor-

relation of −.23 with nine types of 'serial' reactions of varying degrees of difficulty. Four 'discrete' reactions in the second experiment yielded an average correlation of .29 with six serial reactions. Neither of these values was statistically significant.

7. Intercorrelations of speed in 'serial' motor activities in the first three experiments were, respectively: (1) .24; (2) .38; (3) .27 for the white boys and .57 for the negro boys. The results show rather unusual agreement in the three experiments, as far as the white subjects go, despite variations in the number of processes involved and in the homogeneity of the subjects used. The high intercorrelation in the case of the negro boys is interesting and unusual, as is the higher average correlation for this group between Stanford-Binet mental age and the six speed tests.

8. An unreliable negative correlation (r equals $-.25 \pm .126$) was found between median lid reflex time and median psychogalvanic time of 24 college students. Both reactions were elicited by the same stimulus and were measured simultaneously (when both occurred). When the individual paired time values (not medians) were correlated, irrespective of the subject making the reactions, the coefficient became $.06 \pm .04$. This coefficient would seem to indicate that the conditions determining speed in these two reflex arcs tend to vary independently. The fact that the reactions whose time values constituted the items correlated were elicited each time by the *same* stimulus would seem to lend special significance to the result in question.

9. Simple reaction time showed no significant relationship to latency in either the lid reflex or the psychogalvanic reflex. The lack of correlation among these simple types of activity would seem to constitute a strong argument against the view that rate of nerve conduction (or any other single organic factor) conditions speed in diverse motor activities somewhat uniformly.

10. In the second, third and fourth experiments, 'intelligence' tests were given to the subjects and the resulting scores were correlated with the measures of speed. In the second

experiment, Army Alpha scores yielded a correlation of .29 with composite 'speed' scores based upon the ten functions measured. The correlations with simple reaction time were higher than with the more complex (?) serial activities. In the fourth experiment, on a similar group of subjects, no correlation at all was found between Army Alpha and auditory reaction time. Nor was any significant correlation found between Army Alpha and either lid reflex or psychogalvanic reflex latency. In the third experiment, on twelve-year-old white and negro boys, the average correlations of Stanford-Binet mental age with six serial 'speed' tests were .31 and .62, for the whites and negroes respectively. Possible explanations of this unusually high correlation in the case of the negroes were considered.

It would appear from such results as these that 'speed' is not conditioned uniformly in all activities by some unitary factor or set of factors. Correlations between speed measurements tend to be high when the postural preparation and the types of motor activity involved are similar. But the size of the correlations decreases practically to zero where the processes correlated are strikingly unlike with respect to these two major conditions. If these results prove to be generally valid, it is apparent that such expressions as the 'relation of speed to mental ability' have no meaning. One must specify what sort of 'speed' is meant, since different types of speed variables may have little in common. That there would be some correlation might perhaps be expected, since the same organism is involved and since certain general factors are necessarily common to all sorts of reactions. But the degree of relationship may be so slight as to be negligible in the case of reactions necessitating different types of nervous integration.

An outstanding implication of the results of these experiments is that 'speed of reaction' is not the product of a single organic variable, such as rate of nerve conduction. The factors determining rate of response in simple motor processes —perhaps even of reflexes—are undoubtedly much more complex than such a view implies. The precise nature of the conditions operating within the central nervous system to

determine the rate at which a single or a succession of explicit motor acts occur is unknown. But these conditions may reasonably be assumed to be more complex than mere rate of transmission of the nerve impulse in the nerve fibers. The facilitating or inhibiting effects of posture and of previous excitation are probably much more important than rate of conduction, *per se*. In fact, relatively greater speed in transmission might well result in many instances in interference with the integrative processes essential to the production of the motor act, with a consequent relatively slow response. A motor reaction is the resultant of numerous nerve discharges from diverse higher centers which impinge on the 'final common path.' A quick reaction probably is dependent upon a ready synchronization of these effects in some manner, perhaps as regards the rhythm of nervous discharges affecting the lower motor centers.

The outstanding generalization from these experiments and from similar results of other investigators is that the amount of correlation between measurements of speed of reaction tends to vary somewhat directly with the similarity in the postural and effector mechanisms involved in any pair of reactions correlated. Such results are unintelligible on the basis of a view that an individual possesses a constant grade of neural conductivity which operates uniformly to determine speed in all types of activities. On the other hand, such results are perhaps to be expected on the hypothesis that speed in a given act depends upon the integration of nerve impulses from the higher motor centers, which may be differentially affected by variations in posture, in the source and nature of the stimulus, in the effector organ involved and in the pattern of response required in the situation. A type of neural organization which would function rapidly in one situation might well operate slowly in another type of activity, in which the rhythms of discharge from the several higher motor centers presented greater 'difficulty' with respect to the physiological resolution occurring in the excitation of the lower motor centers directly controlling the act.

(Manuscript received March 8, 1933)

References

1. Aikens, H. A., Thorndike, E. L., and Hubbell, E., Correlations among perceptive and associative processes. *Psychol. Rev.*, 1902, **9**, 374–382.
2. Bernstein, E., Quickness and intelligence, *Brit. J. Psychol. Monog. Suppl.*, 1924, No. 7, 1–55.
3. Dodge, R., *Elementary conditions of human variability*, New York: Columbia Univ. Press, 1927.
4. Farnsworth, P. R., Seashore, R. H., and Tinker, M. A., Speed in simple and serial action as related to performance in certain 'intelligence' tests, *Ped. Sem. and J. Genet. Psychol.*, 1927, **34**, 537–551.
5. Hansen, C. F., Serial action as a basic measure of motor capacity, *Psychol. Monog.*, 1922, **31**, No. 140, 320–382.
6. Hart, B., and Spearman, C., Mental tests of dementia, *J. Abnor. Psychol.*, 1914, **9**, 216–264.
7. Kennedy, Margaret, Speed as a personality trait, *J. Soc. Psychol.*, 1930, **1**, 286–299.
8. Lambeth, Martha, and Lanier, L. H., Race differences in speed of reaction, *Ped. Sem. and J. Genet. Psychol.*, 1933, **42**, 255–297.
9. Leedy, John L., The relation of speed of reaction to basal metabolism and blood pressure, Unpublished thesis in Vanderbilt University Library, 1931.
10. McFarland, Ross A., An experimental study of the relationship between speed and mental ability, *J. Gen. Psychol.*, 1930, **3**, 67–97.
11. Pierce, A. W., A study of speed in certain reflex and voluntary reactions, Unpublished thesis in Vanderbilt University Library, 1930.
12. Peak, Helen, and Boring, E. G., The factor of speed in intelligence, *J. Exper. Psychol.*, 1926, **9**, 71–94.
13. Peterson, J., and Lanier, L. H., Studies in the comparative abilities of whites and negroes, *Ment. Meas. Monog.*, 1929, No. 5, 1–156.
14. Rounds, G. H., Is latent time in the Achilles reflex a criterion of speed in mental reactions?, *Arch. Psychol.*, 1928, **15**, No. 95, 1–91.
15. Seashore, R. H., Individual differences in motor skills, *J. Gen. Psychol.*, 1930, **3**, 38–66.
16. Sisk, T. K., The interrelations of speed in simple and complex processes, *George Peabody Coll. Contr. to Educ.*, 1926, No. 23, 1–48.
17. Tinker, Miles A., The significance of speed in test response, *Psychol. Rev.*, 1931, **38**, 450–454.
18. Travis, L. E., and Hunter, T. A., The relationship between 'intelligence' and reflex conduction rate, *J. Exper. Psychol.*, 1928, **11**, 342–354.
19. Travis, L. E., and Young, C. W., The relations of electromyographically measured reflex times in the patellar and Achilles tendon reflexes to certain physical measurements and to intelligence, *J. Gen. Psychol.*, 1930, **3**, 374–400.
20. Whitehorn, J. C., Lundholm, H., and Gardner, G. E., Concerning the alleged correlation of intelligence with knee-jerk reflex time, *J. Exper. Psychol.*, 1930, **13**, 293–295.
21. Wissler, Clark, Correlations of mental and physical tests, *Psychol. Monog.*, 1901, **3**, No. 16, 1–62.

PRIOR ENTRY WITHIN A SINGLE SENSE DEPARTMENT [1]

BY J. GARTON NEEDHAM [2]

One of the classical experiments in the field of attention is the so-called 'complication experiment,' which is presumed to demonstrate the formation in consciousness of an experience which involves a combination or 'complication' of sensations from two different sense departments. As it was devised first, the experiment required an observer to note the moment of simultaneity of two sensations, one visual, the other auditory, and to report this moment in terms of the visual stimulus (a pointer moving across a clock-like scale). The earliest investigators, Wundt,[3] von Tchisch,[4] and Angell and Pierce,[5] found that subjective and objective simultaneity often failed to correspond, and they were perplexed especially by the appearance of negative displacements, which were cases in which the auditory stimulus was reported as having been perceived before it was actually presented. Studies by Geiger [6] and Stevens [7] seemed to clarify this problem, and made it possible for Titchener to summarize the situation in the Law of Prior Entry: "The object of attention comes to consciousness more quickly than the objects which we are not attending to." [8]

The phenomenon of prior entry was subjected to certain criticisms because of the nature of the stimuli which were employed in demonstrating it. James spoke of visual sensa-

[1] From the Princeton Psychological Laboratory.

[2] National Research Council Fellow.

[3] W. Wundt, *Physiologische Psychologie*, i, 1874, 765 ff.; ii, 1880, 67.

[4] W. von Tchisch, *Philos. Stud.*, 2, 1885, 603–634.

[5] J. R. Angell and A. H. Pierce, Experimental research upon the phenomena of attention, *Amer. Jour. Psychol.*, 4, 1892, 528–541.

[6] M. Geiger, Neue Complicationsversuche, *Philos. Stud.*, 18, 1903, 347–436.

[7] H. C. Stevens, A simple complication pendulum, *Amer. Jour. Psychol.*, 15, 1904, 581.

[8] E. B. Titchener, *The Psychology of Feeling and Attention*, 1908, 251.

tions of position and of motion,[9] and Dunlap[10] extended the criticism by showing that the nature of the displacements was in large part a function of certain visual conditions, especially those of eye-movement and fixation. However, Stone carried out an experiment upon the auditory-tactual complication,[11] and on the basis of her data reaffirmed the validity of the law of prior entry.

Throughout the history of the "complication experiment" the Herbartian usage of the term has been reflected. Without exception the systematic discussions have implied a notion of the relative 'effectiveness' of *two* different sense departments, and of certain modifications of this 'effectiveness' by certain attentional factors. Therefore it becomes a matter of considerable interest to inquire concerning the possibility of demonstrating a comparable phenomenon in a situation where the "complication" is between two stimuli which are received through the same sense department.[12]

Rubin reported recently [13] that under an appropriate *Einstellung* the phenomenal sequence of two qualitatively different auditory stimuli might be the reverse of the objective sequence, if the interval between the stimuli was sufficiently brief (of the order of 100 sigma). Until the *Einstellung* which effects this reversal is defined more clearly, the significance of the experiment cannot be evaluated. Meanwhile, the following procedure was devised, and was found to be appropriate to the problem.

The experimental technique consisted of the presentation to an observer of a sequence of five auditory stimuli (hereafter designated as the 'series'), exactly similar in every respect, and appearing at a constant rate. The duration of each stimulus was 0.25 seconds (250 sigma) and the intervals between

[9] W. James, *Principles of Psychology*, 1890, I, 415.

[10] K. Dunlap, The complication experiment and related phenomena, *Psychol. Rev.*, 17, 1910, 157–191.

[11] S. A. Stone, Prior entry in the auditory-tactual complication, *Amer. Jour. Psychol.*, 37, 1926, 284–291.

[12] The writer wishes to acknowledge his gratitude to Professor H. S. Langfeld for the proposal of the problem and for helpful discussion and criticism.

[13] E. Rubin, Beobachtung zur psychologische Akustik, *Zsch. für Psychol.*, 122, 1931, 109–114.

the successive members of the series were likewise 0.25 seconds. During the presentation of the series, a discrete and qualitatively different auditory stimulus was presented ('discrete stimulus'), its duration being 0.25 seconds. The introduction of the discrete stimulus did not interfere with any of the physical conditions of the series.

The appearance of the discrete stimulus with respect to the series was varied haphazardly in successive trials among five equidistant positions. These settings, in their relationship to the series, are indicated in Fig. 1. Following this figure,

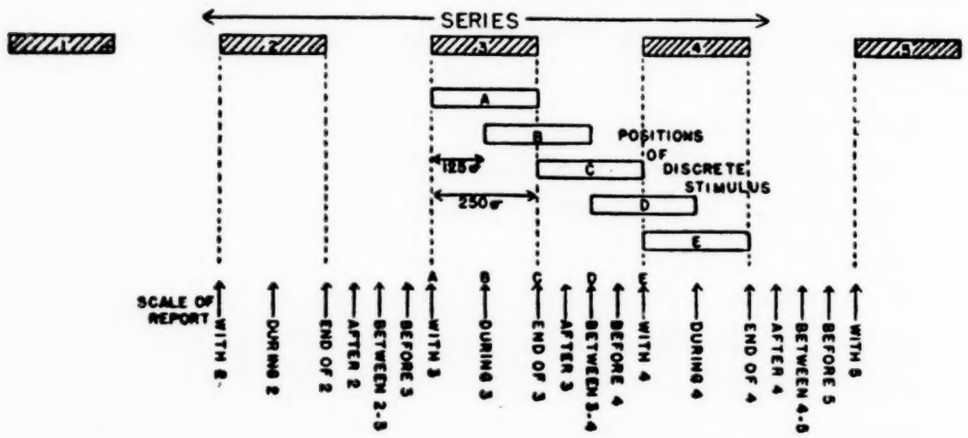


FIG. 1. Representing positions of the discrete stimulus and of reports upon it in reference to the series.

the designation of the five positions of the discrete stimulus as *A*, *B*, *C*, *D*, and *E* will be employed hereafter.

Time controls were provided by an adapted Lewin-Schumann *Zeitsinnapparat;* the timing apparatus was located in and controlled from a room adjacent to that in which the observer sat. The stimuli which were employed were: (1) A relatively pure tone (frequency, 1100 cycles per second; loudness, 60 db above an approximated *RL*) which was heard in an ear-phone at the right ear of the observer. (2) The sounding of a buzzer (aperiodic; loudness not measurable, but estimated to be, at the ear of the observer, approximately equal to that of the tone). The buzzer was attached to a panel extending from the wall, 30 inches from and slightly below the level of the right ear of the observer.

The time relations of the actuation of the buzzer and of

the diaphragm of the ear-phone were controlled carefully. There was a difference in the location of the two sound sources, but experimental evidence was obtained which indicated that this factor did not influence the results.

The task of the observer was to listen to the series and to be prepared to report the point at which the discrete stimulus was first heard; he was instructed to accomplish this by using the series as a scale of reference. The instructions to the observer were verbal rather than written. The emphasis of the instruction was upon the immediate perception of the discrete stimulus. Phrases such as "locate the tone (buzzer) very accurately in the series," "get the buzzer (tone) as soon as it comes," and so on were employed frequently. The observer recorded his reports in columns upon a blank prepared for this purpose. Every possible latitude was permitted the observer in the adoption of categories of report. His task was defined in general terms and he was allowed to carry it out in the manner most convenient to him.

The observers were fourteen members of the undergraduate course in experimental psychology. Each observer gave in a single session 100 reports, 20 upon each of the five positions of the discrete stimulus. This provided a total of 1400 reports, 280 upon each position of the discrete stimulus.

The greater part of the experimentation was carried out with the series composed of tones (1, above), with the buzzer (2) serving as the discrete stimulus. As a check on the physical conditions, a small number of reports was obtained with the buzzer forming the series and the tone as the discrete stimulus. These variations were interspersed haphazardly through the observational period; the observer was warned in advance when such a variation was to be introduced.

In the treatment of the results, the first task was the standardization of the reports which were obtained. An inspection of the records revealed that it was entirely possible to eliminate all individual differences among the observers by a standardization of the terminology of the reports, and thus to arrive at a general scale along which all reports could be

assigned.[14] The reported positions of the discrete stimulus which were represented on this scale are indicated in Fig. 1, which shows also the points on the scale at which the obtained report was the correct one for each of the five positions of the discrete stimulus.

The subsequent treatment of the data involved simply the compilation of the reports obtained for each position of the discrete stimulus; Tables 1 and 2 present these data.[15] From these tables it may be seen that for each position of the discrete stimulus there was a certain number of correct reports (reports of the appearance of the discrete stimulus at its objective position on the scale), and also a number of objectively incorrect reports. In the tables some of the latter appear above and some below the lines which demarcate the region of correct report. These errors may be considered as either negative or positive; they are negative when they involve a report of the discrete stimulus in a position *before* the correct position (above the upper line), positive when the discrete stimulus is reported *after* the correct position (below the lower line).

On the basis of this terminology it is possible to prepare the lower part of Tables 1 and 2, in which the positive and negative errors are summated, and the respective percentages of correct reports and of directed errors are computed. For example, in the case of position *A* in Table 1 (buzzer as discrete stimulus) there are 92 reports of the discrete stimulus in the correct position. Of the remaining reports upon position *A*, 137 are negative (prior to the correct position on the scale of report) and 28 are positive errors (below the correct position). This summary is carried out for each position of

[14] An example will serve to demonstrate that this standardization involved no arbitrary treatment of the data. The following reports were grouped together under the heading 'With 3' (meaning that the discrete stimulus was first heard simultaneously with the first perception of the third member of the series): 'With 3,' 'At 3,' 'Beginning of 3,' 'On 3,' '3.'

The high degree of consistency from one observer to another in the adoption of a system of report validates the use of a prearranged scale of report, based on the present data, in future experimentation.

[15] The total number of reports included in the tables fails to equal the total number of tests made (1,400); this is due to a small and irregular number of instances in which the observer failed to make a report.

TABLE 1

Report	Position of Discrete Stimulus				
	A	*B*	*C*	*D*	*E*
With 2	2				
During 2					
End of 2	2	1			
After 2	28	6	1	1	
Between 2–3	34	7	4		
Before 3	71	25	11	7	
With 3	92	169	57	22	4
During 3	16	40	34	14	3
End of 3	2	9	24	11	2
After 3	10	12	98	77	41
Between 3–4		3	44	99	71
Before 4		1	2	26	65
With 4			2	9	58
During 4				4	13
End of 4					
After 4					
Between 4–5					1

SUMMARY

	Distribution of Reports					Totals	
	A	B	C	D	E	(*A* to *E*, incl.)	(Excl. *C*)
Negative	137	208	107	132	186	770	663
Correct	92	40	24	99	58	313	289
Positive	28	25	146	39	14	252	106
	Distribution in Percent						
Negative	53.5	76.2	38.6	49.0	72.1	57.7	62.7
Correct	35.7	14.6	8.7	36.6	22.5	23.5	27.3
Positive	10.8	9.2	52.7	14.4	5.4	18.8	10.0

Series . . . . . : Tone
Discrete stimulus: Buzzer

the discrete stimulus, and the distributions which are obtained are then summated. Thus, Table 1 shows that in the case where the series consisted of tones, with the report being made upon the appearance of the buzzer, 58 percent of all reports involved a negative error, that is, a report that the buzzer was heard before it actually appeared, while only 19

percent of the reports were classified as positive errors. Furthur inspection of Table 1 reveals the fact that this preponderance of negative over positive errors is present in the case of each position of the discrete stimulus, with one exception. The reports upon the position designated as *C*, the actual position of which may be seen in Fig. 1, do not involve this relationship. Correct judgments are in the minority, and among the errors the positive rather than the negative errors preponderate. Since position *C* seems to represent a special case, the last column of Table 1 presents the distribution of the total reports upon the four positions *A*, *B*, *D*, *E*. The result of this tabulation of the data obviously increases the support for the generalization stated above. Negative errors

TABLE 2

Report	Position of Discrete Stimulus				
	A	*B*	*C*	*D*	*E*
After 2	5	2			
Between 2–3	1				
Before 3	8	1			
With 3	6	4		1	
During 3	4	1		2	2
End of 3					
After 3		1		5	8
Between 3–4				2	5
Before 4					5
With 4				1	4

SUMMARY

	Distribution of Reports					Totals
Negative	14	7		8	20	49
Correct	6	1		2	4	13
Positive	4	1		1		6
	Distribution in Percent					
Negative	58.3	77.8		72.7	83.3	72.0
Correct	25.0	11.1		18.2	16.7	19.1
Positive	16.7	11.1		9.1	0	8.9

Series . . . . . : Buzzer
Discrete stimulus: Tone

comprise 63 percent of the reports, correct reports 27 percent, and positive errors 10 percent.

Table 2 presents the data from the few cases in which the buzzer was employed for the series, with the tone as the discrete stimulus. Five observers, selected at random, gave these reports, under the conditions suggested above. At the time of the experimentation, the anomalous results for position *C* had not been revealed; unfortunately, in the interests of time and convenience, the majority of the reports obtained under these reversed objective conditions were for positions *A* and *E*, with a small number for *B* and *D*, and none for *C*. Because position *C* is not included in the data, the summated distributions in the last column of Table 2 are directly comparable with the final column of Table 1. Such a comparison is entirely favorable; 72 percent of the reports involved negative errors, 19 percent were correct reports, and 9 percent were positive errors. The correspondence of these values with those cited above (with *C* eliminated) seems adequate proof of the absence of any dependence of the nature of the distribution of errors upon the objective conditions, that is, any gross differences between buzzer and tone, in intensity, quality, or position.

One further refinement of the data seems justified. An inspection of the distributions of reports for the individual observers suggested that there were four of these observers whose reports involved relatively less of the tendency towards a marked preponderance of negative errors. Since none of these four observers was among those selected at random for the obtaining of the data presented in Table 2, Table 1 is the only one which is concerned in a division of the records on the basis of individual differences. Table 3 presents the data for the summation of reports into the three categories, and the corresponding percentages, for two arbitrary groups of observers, one of 10, the other of 4, observers.

From Table 3 it may be concluded that for the majority of the observers the preponderance of negative errors is even more marked and consistent than was suggested by Table 1. For the group of 10 observers, excluding the reports upon

TABLE 3

	Position of Discrete Stimulus						
	A	*B*	*C*	*D*	*E*		
Summary—10 Observers							
	Distribution of Reports					Totals	
						(*A* to *E*, incl.)	(Excl. *C*)
Negative	127	153	79	109	161	629	530
Correct	42	27	2	68	17	156	154
Positive	11	14	117	14	1	157	40
	Distribution in Percent						
Negative	70.6	78.9	39.9	57.1	90.0	66.8	73.2
Correct	23.3	13.3	1.0	35.6	9.4	16.6	21.3
Positive	6.1	7.8	59.1	7.3	0.6	16.6	5.5
Summary—4 Observers							
	Distribution of Reports					Totals	
						(*A* to *E*, incl.)	(Excl. *C*)
Negative	10	55	28	23	25	141	113
Correct	50	13	22	31	41	157	135
Positive	17	11	29	25	13	95	66
	Distribution in Percent						
Negative	13.0	69.6	35.4	29.1	31.6	35.9	36.0
Correct	64.9	16.4	27.9	39.2	51.9	39.9	43.0
Positive	22.1	14.0	36.7	31.7	16.5	24.2	21.0

Series . . . . . : Tone
Discrete stimulus: Buzzer

position *C*, negative errors comprise 73 percent of the reports, while only 5 percent of the reports involve positive errors. In the data for the 4 observers, there is little significant preponderance of any one category of reports over another; although the percentage of negative errors exceeds that of positive errors, correct reports give values larger than either.

One point worthy of note concerns the reports assigned to the discrete stimulus in position *C* by the two groups of observers. It appears that the anomalous results noted for this case are provided by the group of 10 observers; it seems to be absent in the data for the 4 observers, as is indicated by the consistency of the totals with and without the reports upon position *C*. That is, the observers who give to the discrete stimulus in positions *A*, *B*, *D* and *E* the greatest and most uniform negative displacement, at the same time give for position *C* a markedly bimodal distribution of reports, with very few correct reports, and a preponderance of positive errors.

From Table 3 it is possible to derive the experimental evidence which sustains most adequately the generalization as to the preponderance of negative errors; when the various anomalous or unfavorable data are withdrawn (last column of Table 3 for 10 observers), the ratio of negative errors to correct reports is 7 : 2, and of negative to positive errors, 15 : 1. By inspection of the data, without any determination of measures of reliability, it seems perfectly consistent to state that under the conditions of this experiment there was a marked tendency to assign to the discrete stimulus a reported position in advance of its actual position with respect to the series; it is obvious that this is directly comparable to the phenomenon of prior entry in the complication experiment.

There are two aspects of the data which must await further experimentation for their explanation. One is the inconsistency of the data for position *C*. Fig. 1 indicates that this position of the discrete stimulus is unique; it is contiguous to both the third and fourth members of the series. The bimodal distribution of the reports for this position suggests that it is so situated that there are two opposed tendencies, one to set it forward, the other to set it back in the series, and this may be due to its unique position. The second point concerns the different degrees of relative preponderance of negative errors among the various positions of the discrete stimulus. The fact that for the comparable positions *A* and *E* the percentage of negative errors is greater for position *E* might suggest that the factors which produce the displacement become more effective as the series proceeds. However, a comparison of the data for positions *B* and *D* fails to confirm this possibility. Both of these matters require further investigation.

The statement that the data presented above demonstrate a case of prior entry may be analysed further in terms of the

actual performance of the observer in the experimental situation. At the close of each observer's service he was requested to state briefly 'how' he went about the task of making the required report. Undergraduate introspections are frequently superficial; however, when the arbitrary classification of the observers into groups of 10 and 4 was made, it was found that the protocols seemed to justify the division. Typical statements may be quoted; from the group of 10 observers, one said, "I get the rhythm of the tones by counting and get the buzzer as soon as it comes," and another, "I beat the rhythm of the series, with my hand, for example, and fit the buzzer into this measure. This I do as the buzzer comes, thus making a unit of the series of six." From the group of 4 observers may be taken such statements as "I listen to the tones and then figure back to get the position of the buzzer," or "I attend to the tones; after the tones have stopped I retrospect to get the position of the buzzer."

Two points appear from the introspective reports. In the first place, it appears that there were two distinct attitudes toward the experimental problem. Either the observer might be actively predisposed toward the buzzer (discrete stimulus), or he might wait until the completion of the series before attempting to estimate the position of the discrete stimulus. Secondly, those observers who adopted the former attitude (which the instructions attempted to inculcate without revealing too much of the purpose of the experiment) made use of secondary aids in 'holding' or 'retaining' the series while they were 'expecting' the discrete stimulus. The method which was employed most frequently was counting, usually verbal, but often supplemented by kinesthetic cues.

These differences are interesting in a theoretical consideration of the concept of attention. It is apparent that in the present experiment the negative errors were favored by 'attention' to the discrete stimulus, and this accords with the findings of Stevens and Geiger. On the other hand, it is impossible to argue that in those cases where this prior entry appeared the series was not attended to; it was necessary that the observer constantly be aware of the series, since it was

essential for the preparation of a report. Thus it seems reasonable to suggest that there were involved two 'levels' of attention, and that the 'set' for the discrete stimulus was suficiently different from the attention to the series to bring about the characteristic results of the present experiment.

As to the general aspects of the phenomenon of prior entry, the results of the present experiment validate a simple restatement of the classical law of prior entry: the stimulus which is more attended to is perceived sooner than that which is relatively less attended to. The fact that this law obtains in the case of characteristically different stimuli presented to the same sense organ signifies that the facilitation by attention of one afferent path as against another is not the explanation of the phenomenon, any more than is the concept of attention as favoring one sense receptor as against another.

Summary

When a discrete stimulus is presented during a series of stimuli of the same sense modality, and is reported by reference to the series, a 'set' toward the discrete stimulus causes the discrete stimulus to be reported as being perceived prior to the moment of its actual occurrence. This displacement is considered to be a case of prior entry within a single sense department.

(Manuscript received April 14, 1934)

FURTHER EXPERIMENTS IN THE STRENGTH OF CONNECTIONS WHERE THE CONNECTION IS PUNISHED OR REWARDED OR NEITHER PUNISHED NOR REWARDED

BY IRVING LORGE, JACK EISENSON, AND BERTRAM EPSTEIN[1]

Teachers College, Columbia University

Lorge and Thorndike ('33) determined differences in strengthening among connections that were left ambiguous or neutral as contrasted with connections that were punished or rewarded. The results showed that "strengthening by the punished occurrences is thus 95 percent, 70 percent, or 90 percent of that by neutral occurrences according as we used first, double, or treble occurrences. It is 41 percent, 44 percent, or 53 percent of that by rewarded occurrences using first, double and treble occurrences."

These facts, as was suggested, may be accounted for without assuming that punishment weakens anything. A connection that is left ambiguous permits the learner to retain a belief that his responses may have been right if for any cause he has such a belief, whereas a punishment does not. Such an uncontradicted belief may well act to some extent as an actual reward would act.

The purpose of this paper is to present further evidence as to the relative potency of reward and of punishment contrasted to ambiguity (neither reward nor punishment) as determined by several additional experiments.

EXPERIMENT 9

The learning task of this experiment was to connect with an

[1] The senior author designed the experiments, supervised the computation, conducted Experiment 9 with Group I, and is responsible for the report and interpretation of the results. Mr. Eisenson conducted Experiment 9 with Group II. Mr. Epstein conducted Experiment 10.

indifferent stimulus word a number from 1 to 5.[2] The directions were:

"This is a learning experiment. I shall read a word to which you will respond immediately with a number from 1 to 5. Sometimes I shall tell you whether you are right or wrong. More often, I shall not tell you whether you are right or wrong, but instead I shall tell you a word so that the time interval after your response shall be filled in the same way as for the responses to which I tell you whether you are right or wrong.

"After every eight responses I shall sometimes give you a shock if your right responses in the preceding eight responses were less than chance, and I shall sometimes give you a token worth one-half cent if your right responses in the preceding eight responses were much better than chance. If your right responses in the preceding eight were just at chance, I shall neither give you a shock nor any tokens.

"The shock will not hurt you, but will be to remind you to better your score in the next trial. There will be five trials with each series.

"Remember, you are to respond with a number just as soon as I say the stimulus word. I shall cash your tokens at the end of each series."

There were forty stimulus words in each set. For thirty-five (Part I) of these words, after the subject had responded with a number from 1 to 5, the experimenter said 'Right' or 'Wrong'; or a pleasant word such as 'adorable, tulip, or cadence'; or an unpleasant word such as 'stench, cuspidor, or sewage'; or an indifferent word such as 'bustle, cortex or endless.'[3]

For the responses to the other five stimulus words, the experimenter said an indifferent word and did nothing, or administered an electric shock, or gave one or two tokens worth half a cent each. These last five stimulus words were always placed so as to be the eighth, or the sixteenth, or the twenty-fourth, etc., word read in a trial. Preceding each trial the thirty-five cards of Part I were thoroughly shuffled so as to equalize the effect of spread of reward (Thorndike, '33a and '33b). It is with the responses to the words of Part I that this paper is concerned.

The subject was seated next to the experimenter, but screened from him so that the experimenter's manipulation was not seen. On one of the hands of the subject were placed the two electrodes of brass gauze. These electrodes were

[2] The indifferent, pleasant and unpleasant words were determined by Dr. Thorndike on the basis of likes and dislikes of words by college students. Through his kindness, the preliminary inventories were made available for this experiment.

[3] See footnote 2.

leads from a Jenkins electro-stimulator, interrupted by a telegraph key on the experimenter's side. The key was depressed to give the shock which is approximately 1.2 milliamperes. The shock was specified as 600 volts administered through a fixed resistance of 500,000 ohms.

After the electrodes had been secured, and after the subject had read the directions, the experimenter said:

"Before we begin this experiment, I am going to give you an electric shock such as will be given during the experiment. This is done to acquaint you with the nature of the shock."

The stimulus word was then read to the subject, and the post-response word was said in a clear and unemotional tone but not as loudly as the stimulus word.

Two groups of subjects participated in this experiment. Group I consisted of twenty paid college undergraduates (all women). The payment consisted of a basal sum for the hour plus whatever bonus was earned. Each subject responded to the words of two sets of stimulus words in one sitting during January 1933. Group II was composed of twenty subjects who were friends of Mr. Jack Eisenson. They received no basal hour payment, but were paid only the bonus money. They responded to the stimulus words of three sets of forty in one sitting during February 1933. This group was more heterogeneous in interests, in ability and education than the college group.

We measure the strengthening effects of the different after-effects upon the connections for which they were announced, by determining the percentage of repetition of the connections with each of the specific after-effects.

Table 1 gives the percentage of repetitions of the same connection in the trial following the one in which a specific after-effect was announced for it. By these determinations, the differences in the potency for such repetition among the various announcements can be ascertained. In Group I, the announcement of 'Right' caused repetition in excess of the percentages for the announcements of 'Wrong,' a pleasant, unpleasant and indifferent word of 22.7, 19.9, 20.9 and 21.3 respectively. The percentage of repetitions attributable to

the announcement of 'Wrong' was 2.8, 1.8 and 1.4 less than the corresponding percentage for pleasant, unpleasant, and indifferent words. The differences among pleasant, unpleasant, and indifferent words were P-U, +1.0, P-I, +1.4, and U-I, +0.4. The differences between the effect of 'Right' and that of the other announcements were reliable; the differences among 'Wrong' and all other announcements need more data to demonstrate their reliability.[4]

TABLE 1

EXPERIMENT 9: THE PERCENT OF REPETITION IN THE NEXT TRIAL OF CONNECTIONS HAVING ATTACHED TO THEM A SPECIFIED AFTER-EFFECT

Group I: Sets A + B

Type of After-effect	Percent of Repetition in Trial 2 for Connections Made in Trial 1	Percent of Repetition for All Other First Occurrences of Connections in the Next Trial	Percent of Repetition of All First Occurrences of Connections in the Next Trial
Saying Right	48.6	44.6	46.0
n	109	202	311
Saying Wrong	23.1	23.4	23.3
n	295	427	722
Saying a Pleasant Word	25.2	26.7	26.1
n	433	795	1228
Saying an Unpleasant Word	24.8	25.2	25.1
n	440	773	1213
Saying an Indifferent Word	22.8	25.9	24.7
n	123	205	398
Group II: Sets A + B + C			
Saying Right	36.4	38.6	37.7
n	176	272	448
Saying Wrong	32.8	28.0	29.8
n	412	654	1066
Saying a Pleasant Word	32.2	27.7	29.4
n	667	1067	1734
Saying an Unpleasant Word	31.9	26.9	28.9
n	688	1025	1693
Saying an Indifferent Word	34.5	26.0	29.4
n	177	265	442

* n throughout this paper refers to the total number of cases upon which the percent is based.

[4] The difference divided by the standard error of the difference is used as the

For Group II, the difference among the percentage of repetition followed by the announcement of 'Right' and that of the other announcements was in every comparison greater for 'Right.' The differences were 7.9, 8.3, 8.8 and 8.3 greater than the corresponding percents for Wrong, P (pleasant word), U (unpleasant word) and I (indifferent word) after-effects, respectively. The announcement of 'Wrong' had repetitions in excess of the percentages of P, U and I of 0.4, 0.9, and 0.5.[5]

The difference between the effect of the announcement of 'Right' and all the other announcements in both groups is large, whereas the difference between 'Wrong' and the announcements of U, P and I are small. Additional evidence may be considered, utilizing the repetitions of connections in the trial following the two preceding trials in which the same connection was followed by the announcement of the same after-effect. Table II reports the percentage of repetition in the next trial of connections having attached to them the same after-effect in two successive trials.

In Group I, the percent of repetitions of connections doubly rewarded was 47.7, 34.5, 38.7 and 33.3 in excess of the percents for double effects of 'Wrong,' P, U and I. All these differences are statistically reliable. The percent of repetition for double effect of 'Wrong' was 13.2, 9.0 and 14.4 less than the percents for P, U and I. The difference P-U was 4.2; P-I, 1.2 and U-I, 5.4.[6]

For Group II, the two successive announcements of 'Right' had percents in excess of all other after-effects of 27.9, 20.4, 24.7 and 24.6 for 'Wrong,' P, U and I, respectively. The effect of two successive 'Wrongs' was 7.5, 3.2 and 3.3 less than the percent of repetitions for P, U and I respectively.[7]

criterion for reliability. The ratio is for R-W, 7.25; R-P, 6.44; R-U, 6.76; and R-I, 5.98. For W-P, 1.39; W-U, 0.90; and W-I, 0.52. For P-U, 0.56; P-I, 0.56, and U-I, 0.16.

[5] The reliabilities of these differences as determined by the ratio, difference/σ difference, are: R-W, 2.95; R-P, 3.32; R-U, 3.52 and R-I, 2.63; W-P, 0.22; W-U, 0.51; W-I, 0.16; P-U, 0.32; P-I, 0.00 and U-I, 0.21.

[6] The ratio, difference/σ difference, for these differences are: R-W, 8.85; R-P, 6.90; R-U, 7.68; R-I, 4.59; W-P, 2.93; W-U, 1.94; W-I, 2.06; P-U, 1.00; P-I, 0.17; and U-P, 0.80.

[7] The ratio, difference/σ difference, are: R-W, 5.64; R-P, 4.44; R-U, 5.35; R-I, 4.03; W-P, 1.97; W-U, 0.83; W-I, 0.60; P-U, 1.28; P-I, 0.80; and U-P, 0.02.

TABLE 2

EXPERIMENT 9: THE PERCENT OF REPETITION IN THE NEXT TRIAL OF CONNECTIONS HAVING ATTACHED TO THEM THE SAME AFTER-EFFECT IN TWO SUCCESSIVE TRIALS

Group I: Sets A + B

Type of After-effect	Percent of Repetitions in Trial 3 for the Same Connections Made in Trial 1 + 2	Percent of Repetitions in Trial 4 for the Same Connections Made in Trial 2 and 3; and in Trial 5 for the Same Connections Made in Trials 3 and 4	Percent of Repetitions in the Next Trial for the Same Connections Made in the Preceding Two Successive Trials
Saying Right	71.7	73.5	72.7
n	53	68	121
Saying Wrong	30.9	20.0	25.0
n	68	80	148
Saying a Pleasant Word	43.5	34.7	38.2
n	108	167	275
Saying an Unpleasant Word	36.1	32.5	34.0
n	108	142	250
Saying an Indifferent Word	39.3	39.5	39.4
n	28	38	66

Group II: Sets A + B + C

Saying Right	68.8	69.9	69.3
n	64	73	137
Saying Wrong	44.4	38.6	41.4
n	133	140	273
Saying a Pleasant Word	51.4	46.7	48.9
n	212	240	452
Saying an Unpleasant Word	44.3	44.8	44.6
n	212	212	424
Saying an Indifferent Word	48.3	40.7	44.7
n	60	54	114

Connections which occurred three trials in succession with the same after-effect are repeated in the next trial with the following percentages in each group:

	Group I	Group II
Saying 'Right'	93.9 (66)	75.0 (76)
Saying 'Wrong'	42.4 (36)	62.8 (86)
Saying a pleasant word	54.5 (88)	69.2 (185)
Saying an unpleasant word	42.4 (59)	72.6 (146)
Saying an indifferent word	80.0 (20)	76.2 (42)

Connections which occurred four trials in succession with the same after-effect are repeated in the next trial with the following percentages of repetitions for each group:

	Group I	Group II
Saying 'Right'	97.1 (35)	90.0 (30)
Saying 'Wrong'	25.0 (8)	69.4 (36)
Saying a pleasant word	82.1 (28)	80.3 (76)
Saying an unpleasant word	61.5 (13)	83.3 (60)
Saying an indifferent word	87.5 (8)	78.3 (23)

In general, the percent of repetition for 'Wrongs' is nearly that for P, U and I. The percent of repetitions for connections rewarded by 'Right' is significantly in excess of all other after-effects. We interpret these results as an indication that rewarding a connection with 'Right' favors its repetition in excess of a definite 'Wrong' and in excess of the ambiguous P, U and I. The difference between 'Wrong' and the ambiguous P, U and I can be explained without assuming that the after-effect of 'Wrong' weakens anything. The differences may be attributed to the fact that the ambiguous after-effects may allow the subject to retain an uncontradicted notion that his response was right, and in those instances cause repetitions such as an announcement of 'Right' would bring about.

We report another experiment in which the learning was more difficult, and the ambiguous after-effect more neutral.

Experiment 10

This experiment was concerned with learning to select a particular part of a nonsense figure each of which was divided into six nonsense divisions. Each page consisted of 49 such figures. Each page was overprinted in red with the experimenter's key. The subject, however, viewed these figures through red glasses which enabled him to see the black outlined figures but not the key. Figure 1 is a reduction of a part of a specimen page.

Each subject was fitted with the red glasses, then shown a specimen page. The directions were then read.

"You see this sheet of pictures? Each one is divided into six parts. When I say 'Go,' you will place a mark 'C' in some one of the six parts of the picture beginning here (showing the

upper left hand corner) and going across the page. It is for you to make the mark in the right part. When you get one right, sometimes I will say 'Right' and sometimes I may say nothing at all. When you get one wrong, sometimes I will say 'Wrong' and sometimes I may say nothing at all. You must place marks on the boxes in time with the machine. Each click means a new box. As soon as you finish one sheet, I will take it away and you will immediately begin the next

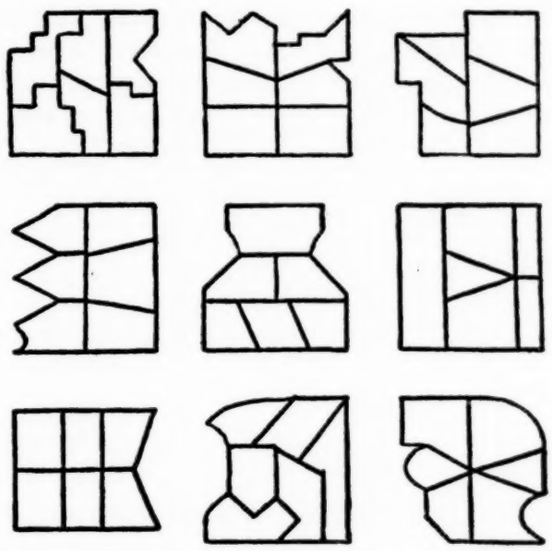


FIG. 1. The type of stimuli presented in Experiment 10. There were seven lines of figures, each line having seven figures making forty-nine in all. The subject marked a segment in each figure moving from left to right across the page every $2\frac{1}{2}$ seconds (Linear reduction one-half).

one. We will see how fast you learned this material in five trials. Ready . . . Go!"

The timer was a motor driven device clicking every $2\frac{1}{2}$ seconds. The subjects experienced no difficulty in keeping time.

The group consisted of 16 college undergraduates (one woman and fifteen men) who marked for five trials, each of four different pages of these figures in one sitting during December 1932.

We can measure the efficacy of saying 'Right' plus a click, of saying 'Wrong' plus a click, or of a click alone by determining the percentage of repetitions in the trial after the one in which connections were influenced by the different after-

effects. Table 3 presents the data for Experiment 10 for all first occurrences of a connection with specific after-effects. The difference between the percent of repetition of connections for the after-effect of 'Right + a click' and 'Wrong + a click' is 10.0; between 'Right + a click' and 'click alone,' 5.3. The percentage of repetition for 'Wrong + a click' is 4.7 less than the percentage for 'click alone.'[8]

TABLE 3

EXPERIMENT 10: THE PERCENT OF REPETITIONS IN THE NEXT TRIAL OF CONNECTIONS HAVING ATTACHED TO THEM A SPECIFIED AFTER-EFFECT

Type of After-effect	Percent of Repetitions in Trial 2 for Connections Made in Trial 1	Percent of Repetitions for All Other First Occurrences of Connections for the Next Trial	Percent of Repetitions for All First Occurrences of Connections in the Next Trial
Saying 'Right' + a click	38.0	29.6	32.6
n	418	751	1169
Saying 'Wrong' + a click	24.3	21.6	22.6
n	2605	4640	7245
Click alone	34.5	22.7	27.3
n	113	176	289

Utilizing the percentage of repetitions for connections in the trial after which the same after-effect was attached to them twice in succession to a connection, we have the results of Table 4. The difference between the percent for 'Right + a click' and 'Wrong + a click' is 21.4; for the difference 'Right + a click' and 'click alone,' 17.7; whereas the difference between the percents for 'Wrong + a click' and 'click alone' is -3.7.[9]

The percent of repetitions in the next trial for connections with the same after-effect three times in succession in the trials immediately preceding are: for saying 'Right,' 72.7 ($n = 139$); for saying 'Wrong,' 38.3 ($n = 345$); for ambiguous 'clicks,' 44.4 ($n = 18$). Four occurrences are followed in the fifth trial with the following percent of repetitions: for saying 'Right,' 84.6 ($n = 65$); for saying 'Wrong,' 43.9 ($n = 82$). The data was too scant for the click-alone.[10]

[8] The ratio, difference/σ difference, is for the 10.0, 6.85; for 5.3, 1.79; for 4.7, 1.76.

[9] The ratio, difference/σ difference, is for the 21.4, 7.11; for the 17.7, 2.78; and for the 3.7, 0.63.

[10] Eight instances.

The difference between the after-effect of 'Right + a click' and 'Wrong + a click' is large; the difference between 'Wrong + a click' and 'click alone' is small. These differences can be explained without assuming that the announcement of 'Wrong' weakened anything, since the ambiguous 'click alone' allows the subject to retain special notions that connections may be right.

TABLE 4

EXPERIMENT 10: THE PERCENT OF REPETITION IN THE NEXT TRIAL OF CONNECTIONS HAVING ATTACHED TO THEM THE SAME AFTER-EFFECT IN THE PRECEDING TWO SUCCESSIVE TRIALS

Type of After-effect	Percent of Repetition in Trial 3 for the Same Connections Made in Trials 1 and 2	Percent of Repetitions in Trial 4 for the Same Connections Made in Trials 2 and 3; and in Trial 5 for the Same Connections Made in Trials 3 and 4	Percent of Repetitions in the Next Trial for the Same Connections Made in the Preceding Two Successive Trials
Saying 'Right' + a click	53.2	50.0	51.5
n	158	172	330
Saying 'Wrong' + a click	32.8	27.8	30.1
n	632	749	1381
Click alone	44.7	20.0	33.8
n	38	30	68

The data from these experiments may be viewed in another way. It is possible to ascertain the excess over *a priori* chance on a per occurrence basis. By *a priori* chance is meant the chance of repeating a connection, if after-effect, special favoritism, or the sheer event of occurrence are considered to have no effect upon the subject. In Experiment 9, this chance is 20.0 (1 in 5). By *per occurrence* is meant the effect of one, two, three or four occurrences of a connection over such chance divided by the number of after-effects producing the after-effect. For instance, the excess over chance of repetitions attributable to the effect of two successive after-effects of the announcement of 'Right' is 52.7 (72.7 − 20). Since two after-effects influenced this value, the value is divided by two to reduce the value to a per occurrence basis, which is 26.3. The weighted [11] average of

[11] Weights on the basis of the frequencies determining each per occurrence value.

the per occurrence values shows the following percents in excess of *a priori* chance for each group:

FOR EXPERIMENT 9

	Group I	Group II
For saying 'Right'	25.5	19.2
For saying 'Wrong'	3.3	10.3
For saying *P*	7.1	11.0
For saying *U*	5.5	10.1
For saying *I*	6.2	10.6

In Experiment 10, the *a priori* chance value is 16.7 (1 in 6). The weighted per occurrence values are:

FOR EXPERIMENT 10

For saying 'Right + a click'	16.5
For saying 'Wrong + a click'	6.1
For a click alone	10.3

The interpretation of the data is in no way changed. 'Right' causes a large percentage of repetition. 'Wrong' does not change the status much less than ambiguity does if allowance is made for the fact that ambiguous after-effects may have in them after-effects of 'Right.' By these comparisons 'Wrong' is 46 percent, 60 percent, and 53 percent of P, U, and I but only 13 percent of 'Right' in Group I; 94 percent, 102 percent, and 97 percent of P, U, and I, but only 54 percent of 'Right' in Group II. In Experiment 10, 'Wrong' is 59 percent of 'click alone,' but only 37 percent of 'Right.'

Punishment, if effective, ought to eliminate the punished connections from an individual's repertory. It does not. Punishment does not overcome the effect of sheer occurrences of connections. The connection has done more harm by occurring than the punishment can offset. In these experiments, ambiguous connections may allow the learner to hold without contradiction the notion that some of the connections are 'Right.' The announcement of 'Wrong' does not. By any reasonable interpretation there ought be a difference as significant as the difference between 'Right' and ambiguity. We have performed at least three experiments in which the difference between 'Right' effects and neutral effects have been large and reliable. Those same experiments, however, show differences between 'Wrong' effects and neutral effects

which are small. The fact that the neutral effects allow the subject to retain special notions concerning the 'Right'-ness of his responses should have favored large differences between 'Wrong' and ambiguity. We do not find such differences. The results of these experiments may be viewed as supporting the earlier findings of Thorndike and of Lorge: in learning, where the subject has no opportunity to react to a punishment by immediately changing his response, that the learning is a function of rewards and not of punishments.

(Manuscript received April 4, 1933)

REFERENCES

LORGE, I. AND THORNDIKE, E. L. ('33), The comparative strengthening of a connection by one or more occurrences of it in cases where the connection was punished and was neither punished nor rewarded, *Journal of Experimental Psychology*, June 1933, **16**, 374–82.

THORNDIKE, E. L. ('33a), A proof of the law of effect, *Science*, **77**, 173–175.

THORNDIKE, E. L. ('33b), An experimental study of rewards, Teachers College Columbia University. Cont. to Ed. no. 580, 72 pp.

THE RÔLE OF KINESTHESIS IN IDEATIONAL MAZE LEARNING [1]

BY WILTON P. CHASE
University of Minnesota

Introduction

Maze learning has usually been considered to be a motor habit. The explanation which has necessarily resulted from this viewpoint has been that it is basically a series of connected kinesthetic patterns. This conclusion has followed from the work of Watson in 1907 and of Carr and Watson in 1908. These workers set up the hypothesis that when a rat learns a maze all that happens is that he becomes conditioned to the series of proprioceptive cues of the running of the maze itself. The stylus maze for humans has been considered as being learned by the same kinesthetic method, by analogy, as for rats.

Recently this viewpoint has been attacked by experimenters in the animal field who have come to regard maze learning by the rats as not being dependent entirely upon kinesthesis.[2] Also, the assumption that human stylus maze learning is a motor or a kinesthetic habit has been attacked.

It seems probable that from the experimental evidence at present available and from a logical analysis that maze learning for animals does not depend upon kinesthesis alone, nor is it a motor habit entirely. For humans, maze learning should be classed probably as an ideational situation, in the same class as nonsense syllables, as contrasted with pure acts of skill definitely involving acquisition of motor habits, such as mirror drawing.[3]

[1] From the psychological laboratory of Dartmouth College. The writer wishes to express his appreciation to Dr. T. F. Karwoski for the helpful advice given by him in this experimentation.

[2] For a review of the literature and a summary of it upon this point, see Heron, W. T., Habit formation and higher mental processes in the rats, *Psychol. Bull.*, 1932, 29, 461–492.

The results of this paper will tend to show that kinesthesis is a relatively unimportant factor in the achievement of learning a maze which is ideational.

Problem

The particular problem to be dealt with in this paper is the influence of kinesthesis in learning a maze of the stylus variety which is ideational in character. In order to work with this problem, two mazes had to be constructed which presented comparable learning situations, except for the variable of kinesthesis which is to be controlled by eliminating its use, as much as it is at present possible to do so, in the learning of one of the mazes. In order to completely and most effectively eliminate the use of kinesthesis, every muscle of the body must be completely paralyzed. This is beyond the present possibilities, and probably entirely unnecessary for all practical purposes. The conclusions which will be arrived at in this experiment, therefore, will be drawn from the statistical results of and interpreted in view of the extent to which the control is exercised over the kinesthetic factor in this experiment.

Apparatus and Technique

A complete description and illustration of the apparatus used in this experiment appears in the *Journal of General Psychology*, 1932, **7**, 485–489.[3]

The apparatus consists of two mazes, a light maze and a stylus maze. Both mazes embodied the same maze pattern. In the light maze, which was set up in a dark room, the subject sees a round point of light which moves in the alleys outlined upon a ground glass surface. The subject sits behind a screen between him and the maze and looks with one eye through an eyepiece which was designed to control eye and head movements and help to keep the subject in a fixed position during the experiment. The experimenter manipulated the maze. The subject gave directions of north, south, east, and west, according to the direction in which he wanted the light moved, and the experimenter moved the maze to follow his directions.

The stylus maze was also especially constructed for this experiment. The stylus fits between grooves to allow the subject to experience the fact that he was touching each side of the alley and to know when he came to a corner, which the construction of this maze allows to become a very definite experience in itself.

---

[3] No detailed historical review will be attempted here, in order to conserve space. References **2**, **3**, **4**, **5**, and **6** have keen included which have something to say on the rôle of kinesthesis or of ideational processes in human maze learning.

The maze pattern used was a modified T pattern which allows for three choices at each corner, either to continue ahead, or to make two turns, each in an opposite direction. Backward moves were eliminated. Usually the subject was interested enough to keep account from which direction he approached the corner, if he didn't, the experimenter told him when he started backwards.

The use of the directions given verbally by the subjects learning the light maze might be criticized as being kinesthetic, and therefore tend to invalidate the conclusions to be drawn. However, they are a necessary part of the procedure and are essential to make the light maze and stylus maze comparable as to the method by which they were learned, which is the counting or verbal method. This is the method which both mazes were constructed to emphasize.

The subjects used in this experiment were psychology students, some being department majors, some being enrolled in the elementary course, and some being enrolled in advance courses. There were 57 in all who satisfied the criterion of learning, 27 in the light maze and 30 in the stylus maze. Thirteen others attempted the mazes, 9 in the light maze and 4 in the stylus maze, but failed to satisfy the criterion of learning through lack of interest to give the necessary concentration to master the method of learning necessary to achieve the criterion of learning established for them at the outset.

The directions given the subject were as uniform as possible for both mazes. The subjects for both mazes were told before starting what the nature of the problem was. To demonstrate this the subject was given a sample section of each maze. Then the subject was told not to hurry, as the time was not being taken into account, rather he should give his consideration to accuracy. The experimenter then placed the subject at the starting point.

Record was kept of trials and errors. The record for the time was omitted, because of the fact that no satisfactory measure of it could be obtained in the light maze since so much of the speed depended upon the experimenter's movements, which would thus be a variable error in trying to measure the subject's time for learning.

An error was counted if a subject entered a blind alley. Backward moves were eliminated by the experimental procedure as indicated above, so were not counted as errors when they were attempted.

The criterion of learning was one perfect trial. Since it was a measure of learning which was desired in this experiment, and one perfect trial represents the threshold of learning, this criterion was chosen as best fitting the purpose of this experiment, rather than a mastery measure, such as three perfect trials.

Results

A. The Light Maze

By referring to part *A* of Table 1, we find that the average number of errors in learning this maze is 350.2 with a standard deviation of 15.18. The standard deviation of the distribution is 78.96 errors. The range is from 214 errors to 504 errors, or 290 errors.

The average number of trials is 21.7 with a standard deviation of .84. The standard deviation of the distribution is 4.35 trials. The range is from 13 trials to 31 trials, or 18 trials.

The average number of errors for each trial, through the 21st trial, with the standard deviation of the average and of the distribution, and the range are also given.

B. The Stylus Maze

By referring to part *B* of Table 1, we find that the average number of errors to learn this maze is 251.2 with a standard deviation of 13.5. The standard deviation of the distribution is 73.98. The range is from 127 errors to 390 errors, or 263 errors.

The average number of trials is 19.53 with a standard deviation of .94. The standard deviation of the distribution is 5.16 trials. The range is from 12 trials to 32 trials, or 20 trials.

The average number of errors for each trial, through the 21st trial, with the standard deviation of the average and of the distribution, and the range are also given.

C. Comparison of the Two Mazes

Part *C* of Table 1 gives the differences between the two mazes. The actual difference between the averages of

TABLE 1

Trial	Errors																					Total Errors	Total Trials
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
A. Light Maze																							
N.....	27	27	27	27	27	27	27	27	27	27	27	27	27	26	25	24	24	24	24	19	16	27	27
Aver.....	40.70	35.19	33.11	28.74	27.52	24.55	21.52	20.07	17.40	16.00	13.81	13.07	11.00	10.08	8.52	7.58	6.63	5.25	4.08	4.00	3.06	350.2	21.7
σ Aver...	1.08	.77	1.18	1.01	.91	.88	.97	.97	.94	1.01	1.01	.97	1.01	1.14	.98	.87	.86	.56	.67	.72	.85	15.18	.84
σ Dist....	5.60	3.98	6.14	5.24	4.74	4.60	5.04	5.03	4.90	5.27	5.23	5.03	5.27	5.80	4.88	4.25	4.22	2.76	3.30	3.14	3.40	78.96	4.35
Range....	31–51	28–48	21–46	18–40	18–39	16–34	12–31	12–28	10–27	7–25	2–24	2–21	0–20	0–29	0–21	0–17	1–19	1–10	0–12	0–10	0–12	214–504	13–31
Diff.....	20	20	25	22	21	18	19	16	17	18	22	19	20	29	21	17	18	9	12	10	12	290	18
B. Stylus Maze																							
N.....	30	30	30	30	30	30	30	30	30	30	30	30	29	28	24	21	21	18	16	13	11	30	30
Aver.....	39.60	31.77	26.37	23.73	19.90	17.13	14.97	12.23	11.43	9.03	7.13	6.73	5.41	5.36	5.67	5.00	4.05	3.61	3.25	2.92	2.45	251.2	19.53
σ Aver...	1.04	.82	.93	1.13	1.06	.96	1.00	.93	1.14	1.21	.83	.75	.74	.81	.82	.74	.67	.61	.75	.61	.58	13.5	.94
σ Dist....	5.68	4.52	5.09	6.19	5.80	5.28	5.50	5.08	6.26	6.64	4.57	4.09	3.99	4.31	4.04	3.41	3.14	2.58	3.01	2.27	1.92	73.98	5.16
Range....	30–51	25–40	15–38	12–39	5–34	7–29	5–29	2–25	2–27	2–30	1–18	0–15	0–15	0–15	0–15	1–12	0–12	0–9	0–8	0–7	0–7	127–390	12–32
Diff.....	21	15	23	27	29	22	24	23	25	28	17	15	15	15	15	11	12	9	8	7	7	263	20
C. Comparison																							
Diff. A–B.	1.10	3.42	6.74	5.01	.762	7.42	6.55	7.84	5.97	6.97	6.68	6.34	5.59	4.72	2.85	2.58	2.58	1.64	.83	1.08	.61	99	2.17
σ Diff....	1.50	1.13	1.50	1.52	1.40	1.30	1.39	1.35	1.48	1.57	1.31	1.22	1.26	1.40	1.28	1.14	1.09	.83	1.00	.94	1.03	20.31	1.26
D/σ Diff..	.73	3.03	4.49	3.30	5.44	5.71	4.71	5.81	4.03	4.44	5.10	5.20	4.44	3.37	2.23	2.26	2.37	1.96	.83	1.15	.59	4.87	1.72

the errors through the 21st trial, for the total errors, and for the total trials of the two mazes are given. The standard deviation and the D/S.D. of Diff. is given for each difference.

Examining these figures we find that there is an actual difference of 99 errors between the averages for the total errors. The D/S.D. of Diff. is 4.87. There is an actual difference of 2.17 trials between the averages for the total trials. The D/S.D. of Diff. is 1.72.

The actual differences for the average errors through the 21st trial vary from .61 error for the 21st trial to 7.84 errors for the 8th trial. The D/S.D. of Diff. also varies from .59 for the 21st trial to 5.81 for the 8th trial.

Discussion of the Results

That the two mazes presented comparable situations is indicated by the averages for the first trial for both mazes Errors made on the first trial in maze learning are made merely by chance, since the subject has no knowledge of what the pattern is like. Therefore, since there is a very unreliable difference of only 1.1 errors between the averages, the two mazes present comparable situations in terms of initial performance by the two groups. If they were different, the chance number of errors would have differed markedly for the two mazes.

An error is an abstraction in itself. However, when averages are compared they show the difficulty of the mazes. Comparing the average total errors for the two mazes, we find that there is a highly reliable difference between them in favor of the stylus maze being easier. Therefore, it seems that although both mazes are learned by the same method, the counting or verbal method, the kinesthetic factor contributes something in the way of ease of learning the stylus maze which is lacking in the light maze due to its control.

An examination of the learning curves for both mazes, as presented in Fig. 1, further bears out this interpretation. Both curves are negatively accelerated. The less negatively accelerated a learning curve is, the more difficult is the learning of the material it represents. The curve for the light maze

shows less negative acceleration than that for the stylus maze, hence, shows that it is the more difficult learning situation. The difference between each of the trials, starting with trial 2 through trial 14, shows an index of D/S.D. of Diff. of 3 or more, which indicates high reliability, therefore, this difference in acceleration is a reliable one. For trial 15 and trial 16, the index for each indicates that we would expect to find the obtained difference 98.6 times in 100; for trial 17,

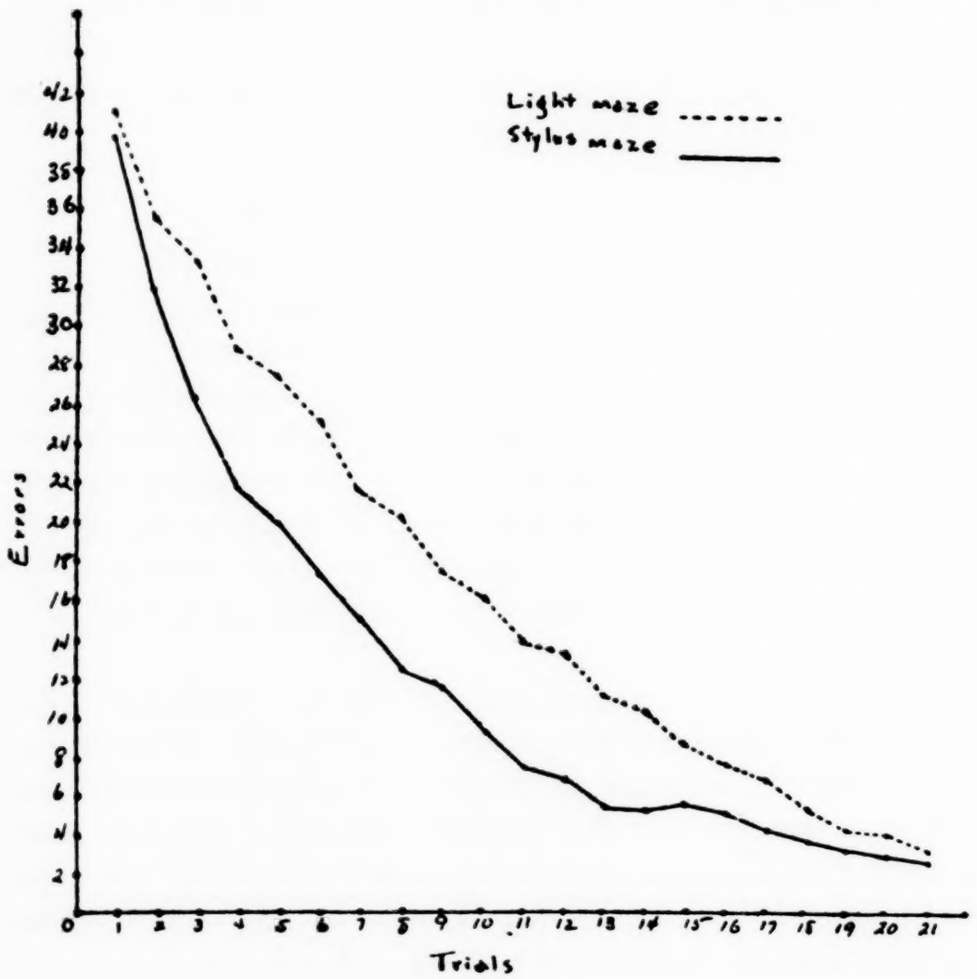


FIG. 1.

98.9 times in 100; for trial 18, 97 times in 100; for trial 19, 79 times in 100; for trial 20, 87 times in 100; and for trial 21, 72 times in 100. From this point there is an insufficient number of cases to accurately carry the curves farther. However, since there is only a small difference between the average number of trials for both mazes, we should expect the curves

to parallel, with possible occasional crossings, with no reliable differences between the averages for individual trials.

The differences between the average number of trials is small, 2.17 trials. The D/S.D. of Diff. is 1.72. This means that 4 times in 100 the difference might be actually reversed. Since the index should be 3 or more to insure high reliability between the averages, the index of 1.72 does not indicate a very reliable difference. The unreliability may be due to an insufficient number of cases. However, in the light of Warden's (**6**) and Husband's (**3**) work which has shown that the difference between the number of trials between the counting method and the motor method is great, it seems probable that no reliable difference between the average number of trials for the two mazes is to be expected, since both mazes are learned by the counting or verbal method.

The kinesthetic factor in the stylus maze contributes to make it easier in terms of errors, as indicated above, but it contributes nothing to the achievement, in terms of length of time or number of trials to learn, this being determined rather by the method of learning. Therefore, we have a situation where kinesthesis at first appears to be an important influence in ideational maze learning, if the errors are examined, but when viewed from the criterion of learning, this advantage appears to be not a significant one in the learning process. In other words, kinesthesis appears to be an artefact as to the apparent advantage which it gives to ideational maze learning by reducing the number of errors. If it were a really influential factor in the situation, the advantage should be carried through and appear to facilitate the achievement by reducing the average number of trials in which the stylus maze is learned. This it does not do.

As to just why the kinesthesis seems to be an advantage in the early trials of the learning process, we have collected some evidence. This will be presented in connection with the next section of this paper.

Special Study of a Selected Group

Since there appears a large significant difference statistically between the error scores of the two mazes, which shows

that the stylus maze is easier in those terms, the question arises as to whether this is due to an actual difference in the mazes, or whether it is due to a difference in the ability of the groups. This might be done by testing both groups on some common problem which correlates highly with the maze learning. However, to find a test which correlates highly with maze learning would be extremely difficult, because as yet none have been discovered. There is but a low correlation between intelligence test scores and maze learning, and correlations between performances on different mazes are usually too low to use as a criterion. Therefore, the method described below was used as the most satisfactory one available at the present time and in the present connection.

A group of six subjects was selected from classes in psychology, three from the beginning course and three from the advance courses. Each subject ran alternate trials on each of the mazes. Three were introduced into the stylus maze first and three into the light maze first. Only a few trials were run in each maze, just enough to allow the subject to gain a comparison between his experiences in the two mazes. Each subject was then asked the following questions,

1. Which maze is the more difficult?

(*a*) Can you give a percentage which will estimate what you think is the difference in the degree of difficulty?

(*b*) What makes the maze easier?

2. Which maze holds your attention better?
3. Which maze holds your interest better?
4. Which maze is less monotonous?
5. What method of learning did you use in the two mazes?
6. For which maze did you have any visual imagery, if at all?

The reports unanimously indicate that the stylus maze seems easier when it is directly compared with the light maze. This helps to explain why it is that there is a difference between the mazes in errors. The use of the hand, involving kinesthetic factors as it does, aids the subject to grasp the spatial relations, directions, and turns more quickly and easily than in the light maze. However, since the pattern is com-

plex and irregular, the subject in the stylus maze can rely neither upon visual imagery nor kinesthesis to successfully achieve the criterion. He must resort in the final analysis to counting the number of squares, and when he does this, the motor advantage of the stylus maze is lost, as is indicated by the number of trials, which are not significantly different from the number in the light maze.

Method of Learning

That each maze is learned by the verbal or counting method is indicated by,

A. Construction of the mazes.
B. The directions given to the subject.
C. The observation of the subject's behavior during the experiment.
D. The subjects' reports.
E. A study of those who failed to satisfy the criterion.
F. The reports of the special cases.

A. Construction of the Mazes

(*a*) *The Visual Maze.*—This maze was learned by the subject, by giving directions for moving the light, *e.g.*, to the east, to the north, to the west, or to the south. When the maze was learned, the subject had a verbal association as follows, east, north, north, west, north, west, west, south, south, etc. This was the method which the construction of the maze imposed upon the subject. He could see but one corner at a time. Following the light, as it moved in the alleys, was merely a device which indicated to him his right and wrong choices, and he was forced to say how the light should be moved. The motor method was eliminated by the construction of the apparatus. If there was any visual method, it was used merely incidently. Husband (**3**, **4**), has shown that the visual method is used but seldom in complicated mazes, and where the counting method is encouraged by the construction of the maze, it is eliminated entirely. The maze pattern used in this maze and the stylus maze is a very complicated one. This helps to insure that the subject shall learn the maze by the counting or verbal method.

(*b*) *The Stylus Maze.*—This maze was learned by the counting of the squares which were traversed by the stylus. When this maze is learned, the subject has a mental organization as follows, (moves) right one, up two, left one, up one, left two, down two, etc. Here again, this was the method which the maze forced upon the subject, due to the construction of the maze in unit squares with a drop into each corner, so that they could be easily counted. The motor method could not be guaranteed to be eliminated by the construction of the maze, because it involves kinesthetic movements. However, the complexity of the pattern and the construction of the pathway to emphasize the counting method, would made it only of incidental importance. Husband (4) has shown that a maze which was constructed to emphasize the counting method was learned by the motor method entirely by only 2 out of 20 subjects. That kinesthesis did contribute something to the learning of the maze is shown by the results obtained in the error scores. However, when we compare the achievement scores in terms of trials we found no reliable differences between the averages for the two mazes. This, as was pointed out, is attributable to the fact that both mazes were learned by the same method.

B. The Directions Given to the Subject

In giving the directions, the experimenter used a practice session where the principle of the mazes was illustrated in which he managed to emphasize the verbal or counting method by going several squares in a row, counting them as he did so. Husband (3) indicates that this is an important factor in obtaining the learning of a maze by a particular method.

C. The Observations of the Subjects' Behavior during the Experiment

This indicated, in the case of the light maze, the use of the verbal or counting method by the nature of their attack upon the problem. They made an attempt to remember which way to move and how many squares to move in each direction. Since they were forced to verbalize the direction, this was relatively easy to observe in all subjects.

In the case of the stylus maze, the subjects could often be heard to count the squares over which they moved. Since, they were not forced to verbalize, this couldn't be observed in all subjects, but only in those who occasionally thought aloud.

D. The Subjects' Reports

When a subject finished the experiment he was casually engaged in conversation about his learning of the maze. It was noticed that in all cases that when they discussed their learning of the maze it was always in terms of the number of squares moved, such as 'two this way,' or 'two east,' or 'one down,' etc. From this fact alone it seems as if we could safely say that all the subjects who successfully satisfied the criterion of learning did learn the maze by the counting or verbal method. Some subjects during the experiment and afterwards said that they tried to visualize the pattern to see what it looked like, but were quite unsuccessful because it was too complex. Most of the subjects were very much surprised afterwards, because they had obtained a very different idea of it from what it actually was.

E. A Study of Those Who Failed to Satisfy the Criterion

What has been said above is further borne out by those who failed to satisfy the criterion of learning. Discussion with those subjects indicated that they failed to pay attention to the counting of the squares as they progressed, because they had no interest in so doing, whether this lack of interest was due to the fact that it seemed too complex and hopeless, whether it was due to some confusion, or whether it was due to an inherent inertia on the part of the subject which was not challenged by the nature of the problem. These failures in both maze situations are directly due to the lack of an attempt to master the mazes by the counting or verbal method. It also explains the selection of those who satisfied the criterion as learning the mazes by this method.

There were more failures for the light maze, because it is a more difficult maze to learn, as is indicated by the reliable difference in the average total errors. It further depended

more directly upon the counting or verbal method, because the motor aid, which made the stylus maze easier, was eliminated. Thus, it is a situation which is harder to find subjects interested to learn.

F. The Reports of the Special Cases

The reports of the selected subjects indicate in the case of three of them that they used the same method of counting in each maze. One reported that he counted in instances in the stylus maze. All reported confusion because of shifting from one maze to the other, and that they didn't really have a chance to concentrate on one to learn the pattern efficiently. Shifting from one maze to the other would, therefore, cause a great deal of interference of one maze for the other in terms of adjustment, interest, attention, etc. However, they experienced no difficulty in transferring the learning of one maze to the other, and in each case learning took place, as was indicated by a smaller number of errors on the last trial which was run as compared with the first trial. These cases merely confirm what has been said before—that to successfully satisfy the criterion of learning the counting method of learning is essential.

The negative indications that visual imagery was of no essential help to these subjects also lend confirmatory evidence that visual imagery is of no importance in learning complex maze patterns.

Summary

1. The problem of the influence of kinesthesis in an ideational maze was attacked in this experiment by an objective technique. Kinesthesis was held constant as a variable, and ideational learning was obtained, both by the nature of the mazes used.

2. A significant difference was found between the average number of errors for the two mazes, indicating that the stylus maze is easier in these terms.

3. The slight difference between the average number of trials for the two mazes was insignificant, indicating that neither maze has the advantage as to the time in which the criterion of learning was achieved.

4. An analysis of the two mazes by a selected group of subjects indicated that the kinesthetic factors in the stylus maze gives it a subjective value as being easier in terms of experience due to the kinesthesis.

Conclusions

In this experiment, all the elements in the learning of the two mazes were the same, except that vision was used in one and kinesthesis in the other, and these were on equal terms in the mazes. The learning in both mazes was ideational. It has shown that although kinesthesis tended to facilitate the learning in the early trials in the stylus maze, this was no real advantage to the achievement of learning the maze, as measured by the criterion of learning. In ideational learning all the sensory modalities are utilized in obtaining the cues for the central mechanism to coördinate in obtaining an insight into the total pattern of the act being learned. More accustomed as we are to using our motor sense in obtaining spatial relations, the use of this factor in the stylus maze enables the individual to gain an initial advantage in this respect. However, since the actual learning depends upon counting the number of squares moved in each maze, the use of the motor sense in the stylus maze fails the individual in obtaining the ideas necessary to achieve the criterion of learning, and it may actually be a hindrance in the final trials if the individual places too much reliance on this factor alone.

These results are in opposition to the conclusions based upon Carr and Watson's original work, and those who have followed Watson in the field of human maze learning, that maze learning is a kinesthetic habit. It favors, rather, the concept that maze learning is a central organization of ideas. That these ideas may be gained through the kinesthetic sense purely, can not be denied, especially where kinesthesis or motor learning is probably the main method of learning. The total relational sense seems to be more important than the building up of successive cues of a kinesthetic nature or otherwise which would lead to a conditioned habit. It probably depends upon some central process depending upon the coördinating of impulses by the cerebral cortex.

The various theories of ideational learning indicate that it is a central process, except possibly the theory of Watson that ideational learning is verbal responses, which means they are conditioned language habits. This neglects the part played by the central mechanism. To limit ideational learning to images is to restrict its scope. Much of our thinking goes beyond imagery and involves symbolization. Many of our ideas come as insight into complete relationship of things. Maze learning as an ideational process shows that insight may come gradually as the relationships are gradually grasped and coördinated in some symbolic manner by the mind.

(Manuscript received April 5, 1933)

References

1. Chase, W. P., A new maze technique, J. of Gen. Psychol., 1932, 7, 485–489.
2. Husband, R. W., Human learning on a four-section, elevated, finger maze, J. Gen. Psychol., 1928, 1, 15–28.
3. Husband, R. W., Analysis of methods in human maze learning, J. Genet. Psychol., 1931, 39, 258–278.
4. Husband, R. W., Comparative behavior on different types of mazes, J. Gen. Psychol., 1931, 5, 234–244.
5. Perrin, F. A. C., An experimental and introspective study of the human learning process in the maze, Psychol. Monog., 1914, 16, pp. 97.
6. Warden, C. J., The relative economy of various modes of attack in the mastering of a stylus maze, J. Exper. Psychol., 1924, 7, 243–288.

MOTIVATION IN LEARNING

II. The Function of Electric Shock for Right and Wrong Responses in Human Subjects

BY KARL F. MUENZINGER

The University of Colorado

This report is a part of a project to study the function of punishment in learning. Its particular purpose was to compare punishment for right responses with that of wrong responses in human subjects as a sequel to a similar study with white rats reported elsewhere (**2**). Instead of devising our own experimental situation we decided to duplicate one used in another laboratory.

Tolman, Hall and Bretnall recently claimed (**4**) to have found, among other things, that the addition of electric shock to a bell signal for the wrong response is less effective (not *more*) than the bell signal alone; and that the addition of electric shock to a bell signal for the right response is neither more nor less effective than the bell signal alone. The subjects in this experiment were college students. At the time it was reported we were carrying on a similar experiment with white rats in which we obtained results that were exactly the reverse of those of the California experiment at least as far as the effect of electric shock was concerned. We found (**2**) that electric shock for the right response was as effective as shock for the wrong response and that either was decidedly more effective (not *less*) than learning without shock. The results with shock for the wrong response are, of course, familiar to all who have worked with animals (**1**, **5**, **6**); we merely supplied evidence to show that shock for the right response has practically the same effect as shock for the wrong response in the visual discrimination habit. In order to gain more insight into this discrepancy between human and animal experiments we decided to repeat and extend the California experiment.

Apparatus

The apparatus of the California experiment "consisted of thirty pairs of holes arranged in an irregular spatial order on a board three feet square.[1] The subject's task was to go through the thirty pairs, punching one hole in each pair with a stylus (analogous to entering the blind or true path section at each of thirty choice points in a maze) and thus to learn for each pair which of the two holes was correct. . . . The holes were three quarters of an inch deep and were backed by uniform brass contact plates. The contact plates of one hole in each of the thirty pairs was connected with an electric circuit so that when the stylus was introduced it rang a buzzer or, in the case of some groups, gave an electric shock plus the ringing of the buzzer. . . . The wiring was so arranged that when the stylus made contact with the plate at the back of the hole it . . . caused an electric shock to pass through the subject by way of the stylus in his one hand and the electrode in his other" (4, pp. 601–603). The shock came from the secondary of an induction coil, and each "subject was asked to regulate the inductorium to give the maximum amount of electric current which he felt he could stand. An attempt to offset the factor of adaptation was made by increasing the strength after each trial" (4, p. 605).

We duplicated this situation with the following exceptions. In the place of holes we used galvanized carriage bolts with round buttons or heads, $\frac{3}{8}$ of an inch in diameter. The subject had to touch these buttons which in certain cases sounded a buzzer and/or gave him a shock. The source of the shock was a power pack with an output of 320 volts; the current varied between 1.8 and 2.2 milliamperes according to the subject's choice. When he completed the circuit he got a single shock and not a succession of shocks as in the case of an induction coil. For the sake of brevity we have called one sequence of holes (those that were starred in Fig. 1 of reference 4) pattern *A*, and the remaining sequence pattern *B*.

Subjects and Procedure

In the California experiment five groups of college students were used under various conditions as shown in table 1. We extended the number of groups to eight for reasons which are apparent from the same table, namely, in order to vary conditions systematically so that certain factors could be compared in strictly comparable groups, a condition not always met in the California experiment. This extension of groups gives us for each situation two sets of comparable groups as shown in Table 2. The number of college students in each of our groups was 26, the number of men varying from 12 to 14 per group. In the California experiment some of the groups were considerably larger than ours, which would no doubt be an advantage.

Each subject was given three trials of forepractice as in the California experiment in which he was asked to touch each button. Then he was told which of the two patterns, *A* or *B*, he was to learn. The instructions, similar to those of the California experiment, were handed to him in typewritten form.

The criterion of learning was one correct trial. In the California experiment the subject was stopped at the 20th trial if the criterion had not been reached. We continued the learning in all cases until the criterion was reached in order to have two more measures, the total number of errors and the number of trials to reach the criterion, with which to compare the groups. It can be seen from Table 2, however, that these last two measures do not give us any results that are materially different from those of the first measure alone.

---

[1] This Punch Board Maze is shown in a cut in 4, p. 602.

TABLE 1

MEANS AND STANDARD DEVIATIONS

Pattern *A* = the pattern of starred buttons or holes had to be learned; *B* = the pattern of non-starred buttons had to be learned. *R* = bell-right or shock-right: the buttons that sounded the buzzer or gave a shock had to be learned. *W* = bell-wrong or shock-wrong: the buttons that sounded the buzzer or gave a shock had to be avoided. *No* = no shock used.

Colorado Experiment										California Experiment		
Description of Groups				Errors for 20 Trials		Total Errors		Number of Trials		No.	Errors for 20 Trials	
No.	Pattern	Bell	Shock	Mean	σ	Mean	σ	Mean	σ		Mean	σ
1	*A*	*R*	*No*	116.1	47.4	122.6	62.7	20.8	9.2	II	103.8	49.7
2	*A*	*R*	*W*	126.3	46.6	142.4	90.0	21.6	10.4			
3	*A*	*R*	*R*	118.0	42.8	124.4	51.3	20.5	7.4	IV	105.7	28.4
4	*B*	*W*	*No*	124.6	54.9	131.7	69.3	19.2	7.3	I	117.2	48.7
5	*B*	*W*	*W*	143.4	52.6	159.0	90.5	21.4	10.2	III	128.8	46.0
6	*B*	*W*	*R*	134.3	53.4	162.7	100.7	22.9	13.1			
7	*A*	*W*	*No*	145.2	46.8	169.5	81.9	25.9	11.3	Ia	108.6	45.8
8	*B*	*R*	*No*	125.5	35.9	137.3	54.4	22.2	7.6			

TABLE 2

COMPARISON OF FACTORS

Critical Ratio = difference between means divided by the probable error of the difference. A positive value indicates that the mean for the first condition is higher than that for the second, or that the first condition is less effective; a negative value indicates the opposite. The entries in italics are for non-comparable groups.

Colorado Experiment					California Experiment	
Conditions Compared	Groups Compared	Critical Ratios			Groups Comp.	Critical Ratios
		Errors for 20 Trials	Total Errors	Trials		Errors for 20 Trials
(*a*) Shock-wrong and shock-right	2–3	1.0	1.4	0.7		
	5–6	0.9	−0.2	−0.7		
	5–3				*III–IV*	*2.99*
(*b*) Shock-wrong and no-shock	2–1	1.2	1.4	0.4		
	5–4	1.9	1.9	1.3	III–I	1.61
(*c*) Shock-right and no-shock	3–1	0.2	0.2	−0.2	IV–II	0.25
	6–4	1.3	2.0	1.8		
	3–7				*IV–Ia*	*−0.35*
(*d*) Bell-wrong and bell-right	7–1	3.3	3.5	2.7	Ia–II	0.61
	4–8	−0.2	−0.5	−2.2		
	4–1				*I–II*	*1.95*
(*e*) Pattern *A* and pattern *B*	1–8	−1.2	−1.4	−0.9		
	7–4	2.2	2.8	3.8	Ia–I	−1.09

Results

In Table 1 we give the means and standard deviations for the various measures, and in table 2 the critical ratios (difference/P.E. of the difference) of the differences between the means. It should be noticed that for each condition to be compared we present in Table 2 two sets of parallel groups, each set containing two strictly comparable groups. For instance, for the first condition, shock-wrong to be compared with shock-right, we have groups 2 and 3 which are alike (pattern *A* and bell-right constant) except for the shock; and we also have groups 5 and 6 (pattern *B* and bell-wrong constant) with the same variation as regards shock. The results from the California experiment are added in both tables in order to afford an easy comparison with our results.

A number of general observations might be made first. Only seven out of the thirty critical ratios are above 2, and only four are above 3. Wherever these larger values appear they are offset by values of the opposite sign in the parallel groups. The parallel groups do not agree very well with each other except in one situation, shock-wrong to be compared with no shock.

Each situation presented in table 2 deserves some special comments:

(*a*) If we compare the effect of electric shock for wrong responses with that for right responses we find practically no difference. Thus if we compare groups 2 and 3 we see that shock-wrong is slightly less effective than shock-right, but the difference is not significant. Or if we compare the parallel groups 5 and 6 we find practically no difference. In the California experiment a greater difference was reported, but it was for non-comparable groups: Their shock-wrong group III learned pattern *B* with bell-wrong while their shock-right group IV learned pattern *A* with bell-right, that is, the two groups varied in all three experimental factors, pattern, bell and shock. In fairness to the California experimenters it should be emphasized, however, that they themselves were in doubt about the comparability of the two groups, at least in regard to the patterns learned, and therefore ran group Ia

with results noted below under (*e*). Perhaps it should also be pointed out here that in the California experiment the only critical ratio that might be considered somewhat significant is precisely this one between non-comparable groups, namely 2.99; all others are below 2. From this fact alone one might immediately conclude that none of the differences in the California experiment can be assigned very definitely to the variation of the experimental factors.

(*b*) If we compare the effect of electric shock for wrong responses with that of no shock at all, we find a consistent, although small and insignificant difference in favor of the no-shock groups. The critical ratios for groups 2 and 1 and the parallel groups 5 and 4 are below 2. In the California experiment a comparison of groups III and I yielded a critical ratio of similar magnitude.

(*c*) If we compare the effect of electric shock for right responses with that of no shock at all, we find no significant difference. Our groups 6 and 4 showed a difference with critical ratios of the same magnitude as the groups compared under (*b*). The parallel groups 3 and 1, however, as well as the California groups IV and II showed practically no difference.

(*d*) If we compare the effect of a bell signal for wrong responses with that for right responses we get such contradictory results that we can safely assume that there is no reliable difference between the two conditions. With our groups 7 and 1 we get a difference in favor of the bell-right condition that is slightly significant. In the California experiment the difference between the similar groups I*a* and II is in the same direction but insignificant. On the other hand, for groups 4 and 8 we get a small difference in the opposite direction.

(*e*) If we compare the difficulty of pattern *A* with that of pattern *B* we again get contradictory results so that we cannot assume a reliable difference between the two situations. With our groups 7 and 4 we get a slightly significant difference in favor of pattern *B*. In the similar California groups I*a* and I there is a slight difference in favor of pattern *A*. In our parallel groups 1 and 8 there is also a slight difference in

favor of pattern *A*. The California experimenters assume on the basis of their results that there is no difference in the ease of learning between the two patterns and that they can therefore compare groups III (shock-wrong; pattern *B;* bell-wrong) and IV (shock-right; pattern *A;* bell-right) as regards the effects of shock for wrong with shock for right responses. Since it is relatively easy to keep the other factors, pattern and bell, constant we preferred to follow the more pedantic procedure of keeping them constant rather than to rely on statistical evidence for the comparability of their variations.

Finally one might ask how the addition of shock-right to bell-right (groups 3 and 1) compares with the addition of shock-wrong (groups 2 and 1); and how the addition of shock-right to bell-wrong (groups 6 and 4) compares with the addition of shock-wrong (groups 5 and 4). Since the critical ratios in all these cases are rather small—none above 2—our conclusion must be that the addition of shock in these cases produces only insignificant differences. For the four conditions named here the critical ratios of the differences between the means of the errors for the first twenty trials are 0.2, 1.2, 1.3, and 1.9 respectively.

As a source of unreliability of all these results we might point out that no attempt was made, neither by us nor by the California experimenters, to have each group contain the same distribution of maze learning ability. It is very likely that the divergence of our parallel groups as shown in Table 2 is a consequence of this neglect. If so, the very existence of such groups might be considered a safeguard against the ready acceptance of spurious differences due to unequal learning ability in otherwise comparable groups.

To summarize these results and evaluations we can state that we found as between shock for wrong responses and no shock at all a consistent, but insignificant difference in favor of no shock and no significant difference between all other conditions that were compared, namely shock for wrong and shock for right responses, shock for right responses and no shock at all, bell for wrong and bell for right responses, and pattern *A* and pattern *B*. In all cases where we had groups

comparable to those of the California experimenters our results agree with theirs, except in one case where the differences between the means were of opposite signs. We definitely disagree with the California experimenters in considering all critical ratios as indicating no significant differences between the means of the groups.

General Interpretation

We can now raise the question: What is the function of punishment by electric shock in human learning?—a question which can best be discussed by reference to the function of such punishment in animals.

In another experiment (**2**) we trained three groups of white rats to learn a black-white discrimination habit. One group was given food for the correct response. The second group was rewarded in the same way and shocked for the wrong response. These two control groups showed the usual differences, the first group acquiring the habit much more slowly than the group that was shocked for the wrong response. The critical ratio for the difference between the means of the errors for the first 100 trials was 14.8. This is merely a confirmation of what has long been known in animal psychology, namely, that if we want to speed up learning we shock an animal for the wrong response. In the present experiment with human subjects such an effect of shock for the wrong response was absent. Must we therefore assume different laws for human and animal learning? [2]

The experimental group of rats was also rewarded with food but shocked for the right response. This group was also decidedly better than the no-shock group, the critical ratio for the difference between the means of errors being 11.7. In the experiment with human subjects the group shocked for the right response did not show any such difference. Again, must we therefore assume different laws for human and animal learning?

It seems to me that such an assumption is unnecessary

[2] The California experimenters do not deny the accelerating effect of shock when given for the wrong response; they merely did not find it in their experiment.

and that the explanation of this discrepancy between human and animal learning is due to a difference in the general motivating factors.

The behavior of a rat is imbedded in a tension produced by 'hunger-food' which can vary considerably in strength as Tolman himself and his students have shown in a series of highly significant experiments (3, pp. 48–67). Since this tension is seldom at its maximum strength it can be increased by the addition of shock.

The behavior of a human subject is imbedded in a tension produced by 'laboratory instructions-perception of goal,' that is, the consent to follow instructions takes the place of the hunger drive in the rat and in place of food as the goal we have the mere achievement of reaching the goal. We consider it superfluous to use food, say in the form of candy, which would not add to the tension which is probably near its maximum strength in the usual laboratory situation. For the same reason we need not expect any increase in tension if we use punishment in the form of electric shock. In other words, we suspect that human subjects in a laboratory situation are so well motivated already that the addition of electric shock has no effect upon their learning efficiency.

If this assumption be correct then we should conclude that adult human subjects do not provide us with a sufficiently significant experimental situation for a study of the accelerating effect of electric shock in learning. On the other hand, it is in experimenting with (children and) animals that we should be able to study that effect. Until we find a more suitable experimental situation for adult human subjects we must consider results from animal experiments as more significant as regards the function of electric shock in learning than results from human experiments.

The California experimenters indorse the report of other investigators that electric shock has a 'disrupting effect,' that it tends 'to hinder learning wherever it is introduced' (4, p. 612). Such an effect probably exists although it is by no means clearly demonstrated by the punchboard maze experiment where the highest critical ratio for the differences be-

tween the means of shock and no-shock groups was never more than 2 as we have shown above. At times our clinical observations of human subjects as well as of white rats have definitely suggested such a disturbing effect even with moderate shocks; with strong shocks it is of course plain as one would expect.

On account of the relative insignificance of all the differences found between comparable groups we think we are justified in not commenting on the three Laws of Learning—Motivation, Emphasis, and Disruption—which the California experimenters suggest on the basis of their results. It may well be that such laws could be substantiated by more suitable experimental set-ups. Our main concern has been with the function of electric shock in the learning of the punchboard maze.

Summary

The experiment with the punchboard maze reported by Tolman, Hall and Bretnall was repeated and extended. As regards the function of electric shock in this situation neither the results of the California experiment nor those of the present experiment showed any significant differences between groups that were and those that were not shocked. It is the writer's opinion that this problem is bound to yield more significant results with animals than with human adults as subjects since the latter are so completely motivated in the usual laboratory situation that the addition of shock cannot produce an accelerating effect upon learning. The writer also admits the existence of a disrupting effect of electric shock although it was by no means demonstrated by the punchboard maze experiment.

(Manuscript received April 17, 1934)

References

1. Hoge, M. A. and Stocking, R. J., A note on the relative value of punishment and reward as motives, *J. Anim. Behav.*, 1912, 2, 43–50.
2. Muenzinger, K. F., Motivation in learning: I. Electric shock for correct response in the visual discrimination habit, *J. Compar. Psychol.*, 1934, 17, 267–277.

3. TOLMAN, E. C., *Purposive Behavior in Animals and Men.* New York: The Century Co., 1932.
4. TOLMAN, E. C., HALL, C. S. AND BRETNALL, E. P., A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption, J. EXPER. PSYCHOL., 1932, **15**, 601–614.
5. VALENTINE, R., The effects of punishment for errors on the maze learning of rats, *J. Compar. Psychol.*, 1930, **10**, 35–53.
6. WARDEN, C. J. AND AYLESWORTH, M., The relative value of reward and punishment in the formation of a visual discrimination habit in the white rat. *J. Compar. Psychol.*, 1927, **7**, 117–127.

A VIEW OF THE CORTICAL INTEGRATIONAL PROCESS THROUGH LIMINAL VISUAL STIMULI

BY M. LUCKIESH AND FRANK K. MOSS

Lighting Research Laboratory, General Electric Company, Nela Park, Cleveland

Since details of objects are seldom, if ever,[1] recognized as a result of the stimulation of a single retinal receptor, the visual sensory process involves some kind of integration of various individual neural stimuli. It is convenient to assume, for the purposes of this discussion rather than as a fact pertaining to the cerebral mechanism, that the integrational process takes place in the occipital cortex. The primary purpose of this study is to present quantitative data pertaining to the character or efficiency of the "cortical" integrational process, as the number of retinal elements involved in seeing a colorless object of a simple geometrical pattern is varied with corresponding changes in the intensity of stimulation by luminous energy. If impulse frequency be considered as the only inherent characteristic of nerve response, it is to be expected that reciprocity relationships will be found between various factors which influence the frequency of the neural currents. The area and the intensity of the stimulus are the adequate factors in this respect.[2] An analysis of the mutual relationships between these two factors provides a basis for studying the "cortical" integrational process.

The analysis is developed from the basis of the relative luminous energy requirements for threshold foveal recognition of objects of various areas and brightness-contrasts with their respective backgrounds. This method of approach is particularly inviting, due to the unusual exactness of available data relating these two factors.[3] However, the interpretation is somewhat complicated by extraneous influences, such as aberration and diffraction phenomena, by the finite characteristics of the retinal mosaic, and by effects due to certain

all-or-none characteristics of visual response. The isolation and appraisal of these extraneous effects permits the formulation of a definite relationship between stimulus energy and number of retinal elements active in the recognition of the test-object. This relationship represents the end-product of a series of reactions including the generation, transmission and utilization of the neural impulses. A glimpse of this cycle of events is presented to establish the quantitative nature of the integrational process.

In general, the present study is presented as a tentative and somewhat empirical expression of a fundamental characteristic of visual perception. As it extends the interpretation of the data beyond the purposes of the original work, it is obvious that the experimental technique of the previous research is not necessarily ideal, although the data are adequate, for an investigation of the present problem.

The Data

Since a detailed description of the experimental technique is available elsewhere,[3] only a brief review of the more important phases of the work are presented here. The subjects viewed the test-object, Fig. 1, binocularly and with the eyes

Fig. 1. The form of the test-object. The interval between the dark rectangles is equal to their width and is one third of their length. The size of the test-object is expressed in minutes of visual angle subtended by the width of the white strip separating the two parallel bars. The test-object was invariably presented to the subjects with the rectangles in either the vertical or horizontal position.

accommodated for distance. The test-object consisted of two parallel bars upon a white background and is described in size by the interval between the bars (which is equal to their width) expressed in minutes visual angle. The contrast between the bars and the background could be readily changed.

The test-object was rotated at a speed above that at which the elements of the pattern are discernible, and the time allowed for recognizing the object was controlled by stopping the rotation for brief and accurately measured intervals. The rotation of the test-object thus provided a confusion-pattern which preceded and followed the brief presentation of the stationary critical stimulus. The subjects were required to report the position of the bars of the test-object as they appeared during the brief presentation in the stationary phase. The visual task involved the resolution of details by foveal vision. The condition of threshold vision was approached by successive reductions in the contrast between test-object and the background. However, the brightness of the visual field was maintained constant, so that retinal adaptation was unaltered. Fifty steps in contrast were established, from 100 per cent to 1 per cent, in an approximate geometric series. The threshold condition was determined by the method of serial groups.[3] The score for each group was the percentage correct answers minus the percentage incorrect answers, and the result from the single series of groups was the value of contrast, interpolated where necessary, corresponding to a 50-percent score. Other definitions of the visual threshold are also justifiable and their significance is suggested later.

The experimental data relating area, contrast and birghtness-level are condensed in Table I from 100,000 observations. They are presented graphically over a limited range in Fig. 2. Lagrange's formula [3] was used to obtain certain interpolations and extrapolations necessary for completing the relationships shown in Fig. 2. It will be noted, in general, that the relative probable errors of the data are considerably less than 0.02 or 2 percent. A precision of this order is evidence of the control of factors extraneous to the purposes of the research. Individual observations were excluded from the averages in such cases as indicated by an application of Chauvenet's Criterion.

Negative Energy

As the white field comprising the background of the test-object may be considered to be indefinite in extent, it is con-

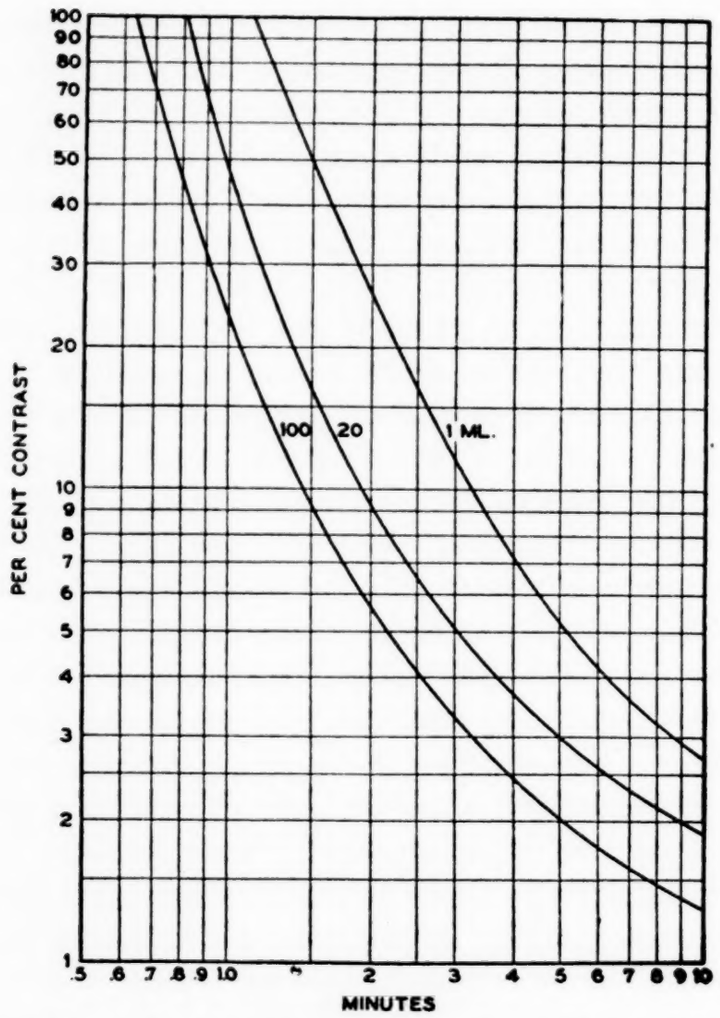


FIG. 2. The relationship between contrast of test-object and its background (expressed as a percentage) and the size of the object (expressed as minutes visual angle subtended by the width of the white strip separating the two bars of the object). The exposure-time was 0.17 second.

venient to express the energy of the stimulus in terms of the energy per unit area subtracted from the visual field by the darker elements of the object. The energy in this case may be termed *negative*, for it represents the amount subtracted when a unit area of the test-object elements replaces an equal area of the white field. The actual unit used is equivalent to the energy subtracted from a visual field, having a brightness of one millilambert, by a square of perfect black subtending an angle of one minute at the eye of the subject when viewed through a pupil one square millimeter in area for 0.170 second. Obviously, this unit of energy is an empirical one,

TABLE I

THE PERCENTAGE BRIGHTNESS-CONTRAST BETWEEN OBJECT AND BACKGROUND REQUIRED FOR THRESHOLD VISIBILITY FOR VARIOUS SIZES OF TEST-OBJECT AND BRIGHTNESSES OF THE VISUAL FIELD. THE EXPOSURE-TIME WAS 0.170 SECOND

The contrast values are geometric means of individual values from nine different subjects of normal or near-normal vision. The probable errors are relative and exclude systematic differences between individual subjects.

Minutes Visual Angle	1 ml.		20 ml.		100 ml.	
	Percent Contrast	Relative Probable Error	Percent Contrast	Relative Probable Error	Percent Contrast	Relative Probable Error
0.80					44.8	0.014
0.99			48.5	0.021		
1.17			30.1	.017	15.2	.017
1.35	60.3	0.016	20.9	.017		
1.63	40.5	.014	14.1	.015	8.13	.013
1.93	27.9	.018	9.96	.015		
2.41	17.2	.021	6.87	.016	4.67	.018
3.20	10.9	.016	4.92	.013		
3.95	7.39	.017	3.71	.012	2.43	.014
5.67	4.43	.014	2.71	.014		
7.98	3.20	.014	2.16	.012	1.50	.010
11.31	2.42	.013	1.80	.013		
16.02	1.95	.009	1.57	.014	1.14	.008

but the actual size of the unit is of no consequence in this study.

In order to simplify the calculation and discussion of contrast, it is assumed that the blackest object used in the experimental work had a contrast of 100 per cent with its background. Actually, the deepest black conveniently obtainable for the test-objects had a brightness equal to 0.021 of that of the white screen, so that the highest contrast was actually 98 percent. Contrast is here defined as the brightness-difference between background and object divided by the brightness of the background and is expressed as a percentage.

STIMULUS ENERGY RELATIONSHIPS

Relative negative energy is computed from the formula, $E = S^2CBP$, where S is the width of the bars of the test-object (and of the distance between them) in minutes visual angle; C, the brightness-contrast expressed as a fraction; B, the

brightness-level of the visual field in millilamberts; and P, the area of the pupil in square millimeters. As pupil measurements were not made in the original investigation, the data of Reeves [4] are used. Since the energy curves for various brightness-levels are quite similar, it is convenient to base the

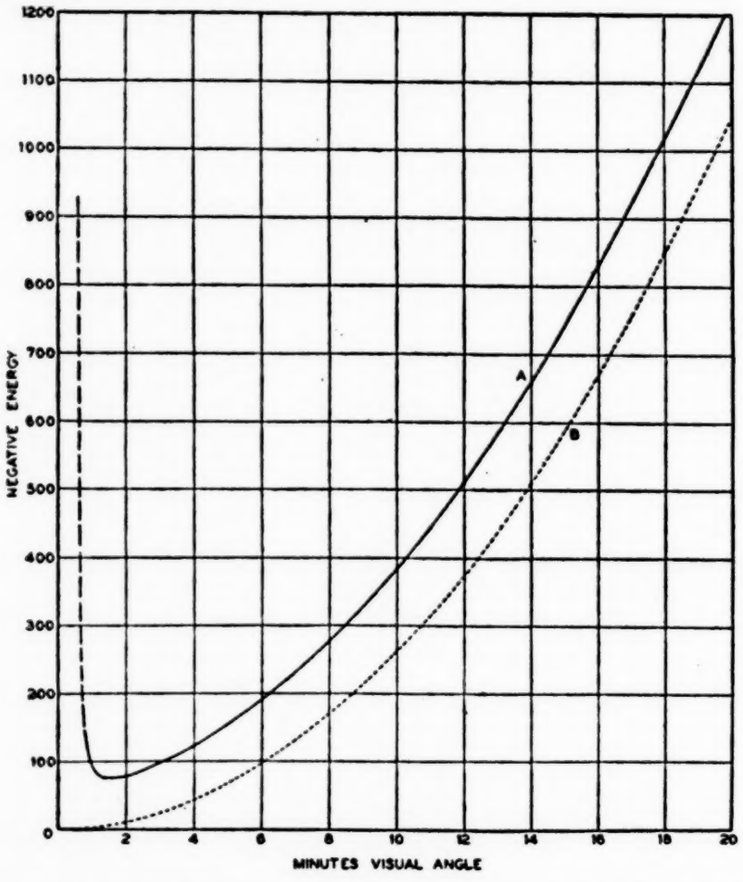


FIG. 3. The relationship between *negative energy* and size of test-object for conditions of threshold visibility is shown by Curve *A*. The solid-line portion of this curve applies to objects of various contrasts of less than 100 percent; the broken-line portion of the same curve applies to objects of contrasts greater than 100 percent. Curve *B* indicates the negative energy required for the perception of the minimum brightness-difference. The curves apply to a brightness-level of 20 millilamberts and an exposure-time of 0.17 second.

discussion upon the data pertaining to a definite brightness of the visual field. For this purpose, a brightness-level of 20 millilamberts is arbitrarily selected.

The socalled negative energy required for threshold vision for various sizes of test-objects and for a constant brightness of the visual field is shown by curves *A* of Figs. 3 and 4. The

solid-line portion of these curves applies to real objects which vary in contrast from a maximum of 100 percent to less than 2 percent. The dotted portion of the curves may be considered to apply to imaginary objects whose contrasts are

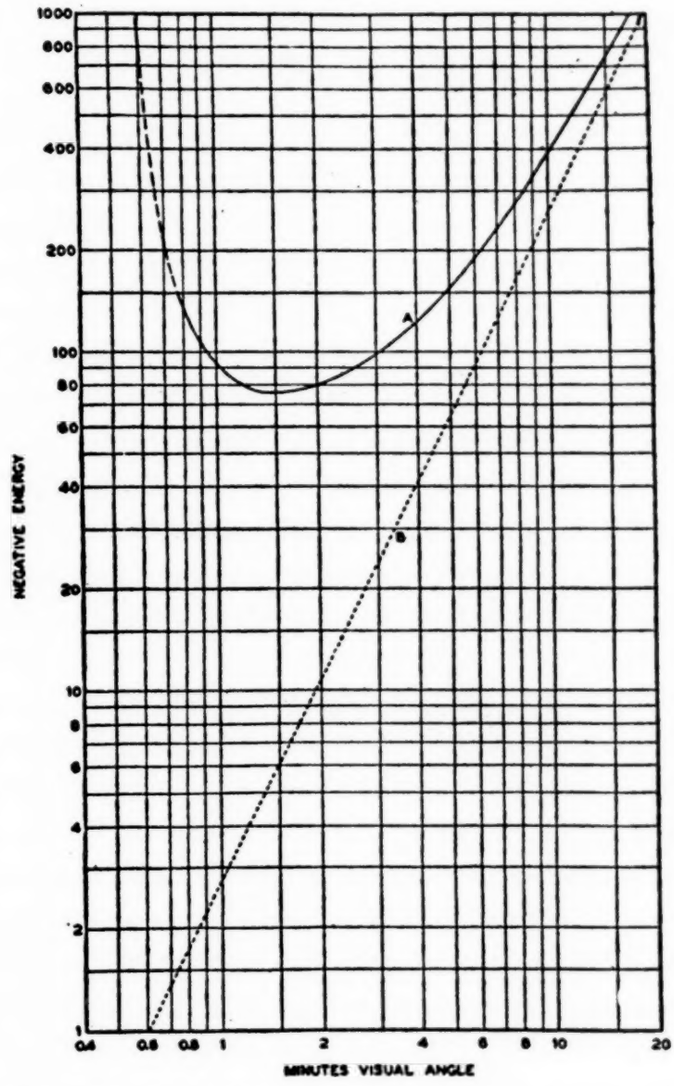


FIG. 4. The same data as Curves *A* and *B* of Fig. 3 plotted on logarithmic coordinates. This method of presentation emphasizes the effects of distortion of the retinal image.

greater than that of black on white and, as such, indicate that the negative energy becomes infinite for an object of finite area.

The relationship between size of object S (minutes visual angle subtended by the width of the sapce between bars of the test-object) and contrast C (percent) agrees closely with an equation of the form:

$$C = a + b(S - d)^{-m},$$

where a and d are the values which C and S will approach asymptotically; and m and b are constants. For a brightness-level of 20 millilamberts and an exposure-time of 0.170 second, this equation becomes:

$$C = 1.28 + 13.97(S - 0.56)^{-1.399}$$

The asymptote of size, indicated as the negative energy approaches infinity, appears to be identical for various brightness-levels and has a computed [3] value of 0.56 minute. As the area of the test-object is increased and the contrast correspondingly decreased, it is obvious that the negative energy required for threshold vision approaches an infinite value. Since the foveal region subtends an angle of approximately 70 minutes, it seems inadvisable to extrapolate the data beyond a size of test-object of 20 minutes. Obviously, when the parallel bars of the test-object are separated by a space of 20 minutes' visual angle, the entire test-object is actually included in a square of 60 minutes on a side.

A conspicuous characteristic of the negative energy curve for a brightness-level of 20 millilamberts is that the energy is a minimum when the size (as defined) of test-object subtends an angle of approximately 1.5 minutes at the eye. The increase in negative energy required for threshold visibility of objects less than 1.5 minutes visual angle is unquestionably due largely to the blurring of the retinal image as a result of diffraction and chromatic and spherical aberration. In addition, the small number of cones included in the image of a small object and their irregular spacing may seriously distort the pattern of the test-object. These factors become insignificant in the case of much larger objects. It might eventually be possible to establish, from complete refractive data, that the increase in negative energy required for objects less than 1.5 minutes is due to these distortions of the image of the object.

The relationships between the extent and the negative energy of the threshold stimulus for various brightnesses of the visual field are shown in Fig. 5. The curves representing this relationship for brightnesses of 1, 20 and 100 millilamberts,

respectively, are based upon extensive experimental data of high accuracy. The data for the other curves were obtained by interpolation. These curves are discontinued at points *a* representing the minimum size of the test-object discernible under conditions of approximately 100-percent contrast. In general, it will be noted that the negative energy required for threshold vision increases with the brightness of the visual field. It will also be noted that the difference between the negative energy required by the smallest discernible test-object (*a*) and the minimum negative energy (*b*) decreases as the brightness-level is decreased. As extrapolation of the experimental data indicates that this difference becomes insignificant for brightness-levels of the order of 0.1 millilambert. That this difference becomes insignificant as the threshold size of test-object increases (as the brightness decreases) is additional evidence that the difference in negative energy at *a* and *b* is a result of the distortion of the retinal image.

The field brightness for each curve in Fig. 5 is constant, but it is seen that the curves are lowered or raised and are always slightly altered for different field brightness. In general, the variation in negative energy with changes in the brightness of the visual field may be considered to represent the degree to which retinal adaptation fails to compensate for changes in field brightness. However, the departure from perfect adaptation is not determinable exactly in terms of negative energy, from these data, since the test-objects of threshold visibility differ in their areal and intensity characteristics; hence definite comparisons are impossible.

Negative-Energy Thresholds

The relationship between extent and brightness-contrast for any liminal visual stimulus indicates a finite asymptotic value of contrasts [3] for each field-brightness. Since the test-object becomes indefinitely large as the asymptote of contrast is approached (Fig. 2), the threshold value of contrast is essentially a datum expressing minimum perceptible brightness-difference. The differences between the asymptotic

values of contrast, which were observed at various field brightnesses, are in accord with the results of other researches which reveal discrepancies in Weber's law.

These data indicate that a definite brightness-contrast must exist between the background and the test-object, re-

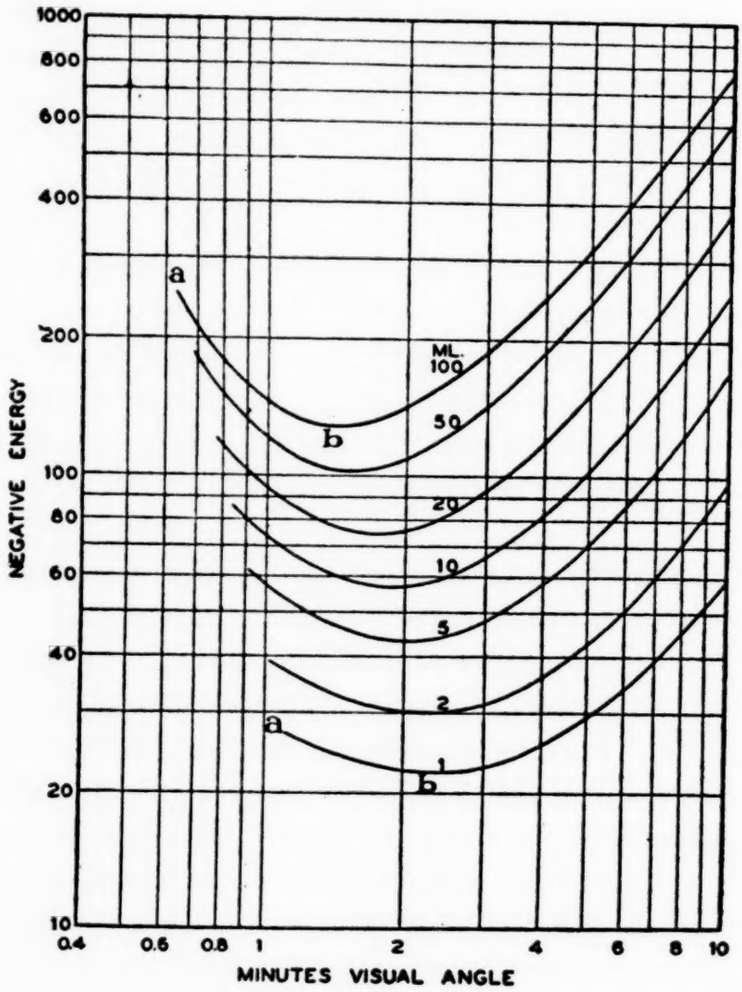


FIG. 5. The relationship between negative energy and size of test-object for various brightnesses of the background. The left extremities, *a*, of the curves represent black test-objects on white backgrounds. The contrasts of the objects with their backgrounds decrease simultaneously with increases in size. The exposure-time is 0.17 second.

gardless of the area of the latter, in order to give rise to a sensation of brightness-difference. Therefore, the asymptotic value of contrast may be considered to represent a functional threshold of brightness-difference which is imposed by an "all-or-none" characteristic of visual response. The broken-

line curves of Figs. 3 and 4 indicate the negative energy required to attain the threshold of perceptible brightness-difference for objects of various areas appearing upon a background illuminated to a brightness of 20 millilamberts. The values of negative energy for curve *B* are computed from the formula, $E = S^2CBP$, where *C* is the threshold of contrast and has a fractional value of 0.0128. For brightness-levels of 1 and 100 millilamberts, the threshold values of contrast are 0.0137 and 0.0086, respectively. These values apply only to the test-object used. Incidentally, Weber's constant is not constant in this case and Weber's law should be used with caution.

These data provide a basis for inquiring into the nature of the processes of summation of neural stimuli in the visual cortex. In the experimental procedure the area and the intensity of the stimulus were varied over wide ranges in such a manner that the test-object was always of threshold visibility. Thus, it was possible to vary both the number of functional retinal elements and their intensity of stimulation without altering the state of retinal adaptation. The assumption of the constancy of retinal adaptation appears to be justifiable, notwithstanding the variations in the brightness of the test-object. At one extreme of size, the bars of the test-object are small and black in comparison with the indefinitely large area of the background. At the other extreme of size, the contrast is low, that is, the brightness-difference between object and background is unimportant and eventually indistinguishable. A comparison of the time required for significant changes to occur in retinal adaptation and the duration of the fixational pauses provides additional proof of the constancy of the retinal adaptation for the various photometric characteristics of the test-object.

The solid-line curves, *A*, of Figs. 3 and 4 represent the total negative energy required for a resolution of the elements of the test-object pattern. A part of this energy is required to bring the image of the object to the threshold of brightness-contrast discrimination, and the remainder provides additional energy evidently necessary for the resolution of the elements

of the test-object. The difference in energy between these thresholds is obviously a function of such factors as the complexity of the pattern and the definition of threshold vision as used in this work. It is conceivable that a more intimate insight into the nature of the integrational process might be obtained from similar analyses involving a series of test-objects of greater geometrical complexity.

Although the difference in negative energy indicated by curves *A* and *B* of Fig. 3 is not the same for all sizes of test-object, it is nearly so in view of the wide range in number of retinal elements involved in the images of the different objects. Actually, the energy-difference increases less than three-fold for a ten-fold increase in the width of the elements of the test-object, or for a hundred-fold increase in area. Such a variation in negative energy indicates (1) that the integrational process for objects of a simple pattern is a highly efficient one and (2) that the efficiency of the process decreases somewhat with increases in the size of the retinal image of the test-object. The influence of brightness-level upon the efficiency of the integrational process is indicated by the data of Table II for arbitrarily selected, though representative, cases. It will be noted that the ratio of the negative-energy difference for the smaller test-object to that for the larger test-object decreases as the field-brightness increases. In general, it may be assumed that the integrational process, as the area of the retinal image is increased, is less efficient at the higher brightness-levels.

TABLE II

THE DIFFERENCE IN NEGATIVE ENERGY REQUIRED TO RECOGNIZE THE PHOTOMETRIC AND GEOMETRIC CHARACTERISTICS OF AN OBJECT FOR BRIGHTNESS-LEVELS OF 1, 20, AND 100 MILLILAMBERTS, RESPECTIVELY

Size of Object	Negative-Energy Difference		
	1 ml.	20 ml.	100 ml.
For Minimum Energy*.....	20	68	117
20 Minutes.....	31	179	553

* The reasons for selecting this point are obvious in Fig. 3.

Conclusion

The assumption is made that the difference in negative energy between curves *A* and *B* of Fig. 3 represents the amount of negative energy required for recognition of the geometrical characteristics of the test-object. The variation in this difference, as the number of receptors involved in the retinal image is altered, is regarded as a measure of the efficiency of the "cortical" integrational process. For test-objects of fairly elementary pattern, the experimental results indicate that the integrational process is a rather efficient one, particularly at the lower brightness-levels. It may also be concluded (1) that the efficiency of the integrational process decreases somewhat as the area of the stimulus increases and (2) that the efficiency decreases as the brightness of the visual field increases. It is emphasized that these conclusions do not involve a consideration of the specific phenomena of color-vision.

The authors wish to express their appreciation of the helpful criticisms of Dr. E. Q. Adams in the preparation of this paper.

(Manuscript received April 5, 1934)

References

1. H. Hartridge, The chromatic aberration and resolving power of the eye, *J. of Physiol.*, **52**, 1918, 175.
2. J. H. Parsons, An introduction to the study of color vision, 1915.
3. P. W. Cobb and Frank K. Moss, The relation between extent and contrast in the liminal stimulus for vision, Jour. Exp. Psy., **10**, 1028, 350.
   P. W. Cobb and Frank K. Moss, The four variables of the visual threshold, *Jour. Fran. Inst.*, **205**, 1928, 831.
   M. Luckiesh and Frank K. Moss, SEEING—A partnership of lighting and vision, Williams and Wilkins Company, Baltimore, 1931.
4. P. Reeves, Rate of pupillary dilation and contraction, *Psy. Rev.*, **25**, 1918, 330.

THE NATURE OF DISTANCE VISION

(As Revealed by a Study of After-Imagery)

BY O. O. NORRIS

Michigan State Normal College

This study was originally prompted, and later guided, by a regard for the scientific postulate of natural or physical continuity. This postulate implies the concept that all designable reality is physical or natural and that therefore, for our purpose here, vision—even after-imagery—must be interpretable in physical-science terms. In this paper, only the matter of perspective will thus be considered. But the conclusions reached may well encourage the conception that the brightness-color phase of vision—its 'qualitative' aspects—must also be amenable to interpretation in physical-science, space-time, processal, structural terms.

The immediate prompting to the study was a flat denial by the late George Herbert Mead of Chicago, in 1922, of an inferential suggestion of the writer, that after-imagery exhibits a perspective, inverse to that of normal, direct, or external vision. The writer was already familiar with the treatment of vision presented in the various textbooks of physics; and with Berkeley's philosophical essays on vision (1), and with the reports of surgical cases that Berkeley and the later editor of his works employed to illustrate and validate his theory (1, 444–448). Mead was at the time lecturing on the nature of the physical, and was holding that we use our tactual and manipular experience to verify and supplement our visual experiences, and thus to distinguish veridical visual objects from illusions, hallucinations, and after-images. But he balked when I suggested that after-imagery had a perspective that was just as truly amenable to physical-geometrical treatment as was that of direct vision. My own logical sense would not easily allow me thus to cut loose from the postulate

of natural continuity, which has had such a steadying and fruitful influence in the history of scientific endeavor; but pressure of other interests long delayed a direct attack upon the problem thus raised.

The problem of the experiment here to be described, then, was to validate—or invalidate—an inference or hypothesis; namely, that after-imagery exhibits a realistic, physical-geometrical perspective, and that this is the inverse of that of original or direct vision. A confirmation of the hypothesis, however, has further implications, which will be considered later.

During the prosecution of the experiment, the writer ran through the available literature of the subject, to see whether his particular point had already been treated. Occasionally he ran across mention of 'Emmert's Law' as having relation to the subject, but did not have at hand the article originally stating it. G. W. Allport restates it in the words, "that the size of the AI varies proportionately with the distance of the projection ground from the eye" (**2**, 111). Here I balked at the expression, 'projection ground,' which seemed to imply a notion that the eye somehow 'projects' its after-images—an expression that seems to savor of belief in the 'evil eye.' I had often reacted in a similar way to use of the word in attempted descriptions of the way in which retinal sensations are referred to external objects, in the processes of distance vision. So here was an added incentive for pursuing the subject. The expression was probably suggested by the name and function of the 'projection lantern'; but the cases of eye and lantern are not at all analogous.

Finally, on appeal to Dr. Norman R. F. Maier, of the University of Michigan, I was furnished with a transcript of the original statement of Emmert's Law: *Die lineare Grösse eines Nachbildes ist . . . gleich der linearen Grösse des Objects, multipliziert mit der Entfernung, aus welcher das Nachbild betrachtet wird, d. h. multipliziert mit der Entfernung aus welcher das Object betrachtet würde, das Ganze dividiert durch die einfache Entfernung* (**3**, 446–447). Emmert makes a correlative statement regarding the areas of AIs.

These statements of the case are confirmed by my own findings; but they apply only to AIs when 'seen' upon a ground. Moreover, the significance of the fact stated, for an understanding of how we learn or attain to adequate distance vision, still remains to be stated. It is this aspect of the problem that makes a brief description of my own attack upon the problem seem most worth while.

The experimenter himself served as subject. The Psychology Department of the University of North Carolina kindly placed at his disposal all the equipment required for the experiment. This consisted of the following materials: a well-lighted room, a large and excellent dark room (8 × 54 ft.), a Stoelting image board, a Nystrom graph chart (ruled into inch squares), tables, clamps, yard stick, meter stick, square, compass, colored papers, light bulbs, switch cords, etc.

Prior to the experiment as finally prosecuted, a conceptual analysis and certain provisional tests were made, to determine the limits within which the experiment should be carried on, and the course that it should follow.

First, it was advisable to ascertain whether the AI exhibited any rotation in relation to the stimulus object. For this purpose, a right-triangular object, 1 × 1½ in. was used. (Any form of object would have served here, significantly marked or notched on one side or corner.) No sort of rotation was discernible. Moreover, when, after fixation upon the object, the head was tilted to right or left, the AI appeared tilted in the same direction, and at a corresponding angle. So also, if, with the object in a vertical position, fixation took place with the head tilted to one side, then, when the head was righted to vertical, the AI was tilted in the direction in which the head was moved. Thereafter, this object was discarded.

Second, it seemed advisable to determine whether brightness and color AIs varied differently in magnitude when the distance of the 'projection ground' or secondary reflector was varied. (As we see by means of light refracted or reflected from an object as primary reflector, it would seem that 'secondary reflector,' or merely 'ground' or 'mat,' would be pref-

erable to 'projection ground' as a designation of the ground upon which the AI is seen. In this way the needed distinction between an AI appearing upon a ground or mat and one appearing without a ground will be clearer.) When it was found that there was no discernible difference, whether the object was red or black, objects of various colors were used, indiscriminately. The general conceptual background had, of course, suggested negative results for both of these inquiries; still, it did not seem well to accept a merely inferential answer to them.

Third, there was need of a basic reference or constant, in relation to which measured deviations of objects and distances could be made and their consequences noted. For this, the line of vision was quite naturally adopted; it was kept level or horizontal throughout, from the eye to the center of the stimulus object, and to the point of fixation upon the AI ground.

Fourth, this preliminary scouting indicated that there were different sorts of AI situations, each of which should be tried out in sufficient detail so that conclusions might be certain. Thus there was possible a gross type of situation in which a secondary reflector functioned, and one in which none was used; *i.e.*, a type in which AIs were seen when the eyes were being stimulated by light from a ground, and one in which the eyes were not being thus stimulated. Further analysis showed that each of these gross types was differentiable into two types crucial for the purpose in hand. Thus the experiment resolved itself into a testing out of AI experience under four different types of situation, which will hereafter be designated and described in order, as Types I, II, III, and IV. Obviously, also, there are possible several cases under each type.

Measurement was involved in several ways: measurement of the dimensions of the stimulus object; of the distances of both the object and the image-ground from the eye; of the angles subtended by both object and ground with the line of vision; and of the AIs. It is obvious that one could not apply the ruler to these last, especially when they were be-

yond arm's length. Here measurement could be only indirect and relative, as the AIs appeared to coincide with one or another number of measured squares on a quadrille-ruled ground, or as their horizontal or their vertical axes coincided with known lengths of the ruled lines on the ground. And in Types III and IV, measurement had to be made by mere subjective estimation. Object and ground—when a ground was used—were, when possible, kept one directly behind the other, the one in front being removed by the hand in order to expose the other.

Type I.—Both axes of the two reflectors were kept perpendicular to the line of vision. The distance from the eye to the primary reflector was kept constant, and that of the secondary reflector was varied, by multiples of this distance. For primary reflector, a one-inch square was used; for secondary reflector, a Nystrom graph chart.

In the preliminary skirmishing, an object ½ inch square had been used, a mat ruled into ½-inch squares, and distances ranging from 6 inches to 3 ft. In the final experiment, the object was kept at a distance of 18 inches, and the distance of the ground was varied, through multiples of this distance; namely, 1, 2, 3, 4, 6, 8, and ½ times, or at distances of 18 in, 3 ft, 54 in, 6 ft, 9 ft, 12 ft, and 9 in. As nearly as could be discerned, the area of the AIs varied directly as the square of the distance of the ground, as compared with that of the object; being, in order, 1, 4, 9, 16, 36, 64, and ¼ sq in. These observations were carefully checked by use of the corresponding metrical distances—50 cm, 1 m, 150 cm, 2 m, 3 m, 4 m, and 25 cm. The same magnitudinal relations obtained here.

Thus my inferential contention with Professor Mead was verified. So, too, was Emmert's Law. But my own contention, as related to the abstract geometry of after-imagery perspective, could now be stated in a more precise way. *The AI of an object, when seen upon a ground vertical to the line of vision, varies in area directly with the square of the distance of the ground in relation to that of the stimulus object.*

Type II.—The distances from the eye to the points of fixation on the primary and secondary reflectors were kept

constant—18 in for each, while the angles that the horizontal and the vertical axes of these subtended with the line of vision, were varied, singly and concomitantly. Here object and ground could not be set in the same horizontal line of vision; the head and eyes had to be rotated through an angle from the one to the other.

For stimulus objects, a 1-inch square, a 1 × 2-inch rectangular oblong, and a 2-inch disk were used; for ground, a stiff card, 9 × 12 in, ruled into ½-inch squares, with alternately light and heavy lines. The angles chosen for the purpose were those of 45° and 22½°. These were chosen simply because of their convenience, as being easily-controlled fractions of the vertical angle, 90°. Only two out of several cases tried out need be presented here.

Case 1.—The object was set vertically to the line of vision. The mat was set with its horizontal axis perpendicular to the line of vision, but with its vertical axis at angles of 45° and 22½° from this line, and its position further rotated, so that the image might be seen upon now the upper and now its lower face.

A. At 45°, each object gave an AI of the same dimension as its own, on its horizontal axis, but about 50 percent longer in the direction of the line of vision—the oblong having been set so that its long axis lay in the direction of this line.

B. At 22½°, each object gave an AI of the same dimension as its own on its horizontal axis, but about 150 percent longer in the direction of the line of vision. Here the farther end of the AI was noticeably wider than the nearer, so that the AI of the disk was actually an oval. The distinction was obviously not so great in *A*. Results for *A* and *B* were the same, respectively, whether the AI was seen on the upper or the lower face of the mat.

Case 2.—The object was kept vertical to the line of vision, as in Case 1; but the vertical axis of the mat was now kept perpendicular to the line of vision, while the near end of its horizontal axis was rotated through angles of 45° and 22½° from the line of vision, both to right and to left, as in Case 1 the vertical axis had been rotated to show AIs both above

and below the line of vision. In *A* and *B* of Case 2, the dimensions of the AIs corresponded with those of Case 1, though their positions were correspondingly changed.

Type III.—After fixation upon the object, the eyes were closed or blindfolded. Here AIs were still 'seen,' but at no specifiable distance from the eyes. They were of the same form and in the same position, but appeared to be smaller than the object. The object was notched on one side, to determine whether by any chance the image, as representing an affected patch of the retina, would appear in inverse relation to the object. No such inversion appeared.

Type IV.—Here an excellent dark room was used. The other apparatus consisted of a closed, opaque wooden box, roughly 5 × 5 × 10 in, containing a 25-watt light-bulb, and with a 1-inch square aperture centered in one end. Two opaque partitions were placed between the bulb and this 'window,' each with a square aperture centered upon the central longitudinal axis of the box. The one next to the window of the box, with an aperture $\frac{3}{4} \times \frac{3}{4}$ in, was set $1\frac{1}{2}$ in back of the window; the other, with an aperture $\frac{1}{4} \times \frac{1}{4}$ in, was set $2\frac{1}{2}$ in back of the first. These partitions, with their apertures and arrangement, were calculated to secure that the light from the bulb should come more nearly from a point. The bulb was propped up so that the filament would be opposite the center of the window. The window was covered with a piece of white bond paper, to obviate glare and excessive lighting of the room. Besides, it was this window, not the light itself, of which an AI was desired. The current could be turned on and off by the experimenter, from any desired distance and position from the box.

With the room darkened and the current turned on, fixation upon the window of the box took place, and then the light was turned off. An AI was 'seen,' apparently and indefinitely 'out there,' as in Type III, but with less apparent diminution in size. Closing and opening of the eyes did not interrupt or in any way affect the image. Holding a mat before the eyes had no more effect than in Type III. No 'act of will' could in any way change the size of the image

or make it conform to that known for any imagined distance of the mat or any other ground. Here, too, there was no inversion of the image.

Thus an AI seen without a ground has no perspective, and hence should be sharply distinguished from one seen with a ground. But in neither Type III nor Type IV could any distinction be made between the AI as 'seen' and any sort of affection of the retina. The retinal sensation, as in original vision, seemed to be the object itself from which the light had come, save in Type III, where the AI was of complementary color.

We can now return to the subject of AIs seen upon a ground, as having a perspective, and explain this perspective in physical-geometrical terms; and thence we can proceed to an explanation of how we learn to see things external to us, as being external. Preparatory to this, however, I shall make some restatements of the facts referred to by Emmert's Law.

AIs when seen upon a ground have a determinate perspective of their own, and this is precisely inverse to that of normal distance vision. In after-imagery seen in the same line of vision as the stimulus object, linear dimensions vary directly with the distance of the ground from the eye; while in original vision, these dimensions vary inversely with distance from the eye. The area of the AI varies directly with the square of the distance, while the area of the transverse plane of an object of external vision varies inversely with the square of the distance. An AI occupies a constant proportion of the visual field at whatever distance, a transverse section of which field itself increases directly with the square of the distance, just as it does in ordinary distance vision; while in original vision, an object at different distances occupies a varying portion of the visual field, and this portion of the field is inversely proportional to the square of the distance of the object from the eye.

One significance of these facts, as thus stated, is that all cases of AIs seen upon a ground are cases of a cone or pyramid of perspective cut by a plane through the line of vision at the given distances and at the given angles with this line.

Thus in Type I this perspective was cut by planes vertical to the line of vision; in Type II, by planes at angles of 45° and 22½°. But this cone or pyramid of perspective is not a merely abstract or 'pure' geometrical form. It is describable as a physical, dynamic structure, further activated by—or as—light; it manifests itself as a cone or pyramid of light.

Nor is this the whole of the matter. The light rays constituting this figure converge in a nodal point, back of the crystalline lens of the eye. Thence they continue to the retina; and thus, in cases of Type I, they constitute, inside the eye, an inverse cone or pyramid, with angles homologous with those of the external light-figure—if we disregard the curvature of the retina and the influences of the several refractive ocular media through which the light passed before its convergence at the nodal point. Thus in any case of after-image of Type I, there are two cones or pyramids of light, apex to apex, both with one straight line as their altitudinal axes, and thus with their bases parallel and precisely inverse to each other, in form. This last aspect of the case is a well-known fact, and has given rise to the problem of why we do not see our objects inverted. But this now becomes a false problem. If its authors had distinctly recognized that there are two cones or pyramids of light involved in every case of after-imagery seen upon a ground, and of external vision as well—both figures distinctly physical as well as geometrical, this historical problem, while raised, might have been recognized for the false problem that it is.

I have indicated that, in my opinion, distance vision represents an acquired or learned, not a native or instinctive, mode of behavior. This learning took place so early in our own development that we cannot recall it. In fact, it starts at so early an age that the infant has no background for recognizing his experiences, and no speech repertory for designating or communicating about them. If he had, he would probably tell us that his first visual objects seemed actually to touch the back of his eyes. But cases of congenital blindness resulting from cataracts, which have been relieved by surgery in later years, should give us a really objective cue

for understanding how the normal infant attains to external vision. Two apparently well-observed and well-reported cases may be cited here, each reported by the operating surgeon.

Both were boys, one between thirteen and fourteen years of age, the other nine. Both had long been able to distinguish day from night, and one color from another; but neither was able to distinguish the shapes of objects or their distance. The report of the first case refers to the boy's experiences after only one eye had been operated upon, about a year before the second was relieved.

> When he first saw, he was so far from making any judgment about distances, that he thought all objects whatever touched his eyes (as he express'd it) as what he felt did his skin; and thought no object so agreeable as those which were smooth and regular, though he could form no judgment of their shape, or guess what it was in any object that was pleasing to him: he knew not the shape of anything, nor any one thing from another, however different in shape or magnitude (1, 444–446).

The report of the second case, made over a century later, quite agrees with this, regarding the subject's judgments of both form and distance. The report does not mention the quite equally important matter of comprehension of magnitude. "The case was that of a fine and most intelligent boy." He had already learned the difference between a cube and a sphere, from manipulation, before the operation on his eyes.

> After keeping him in a dark room for a few days (sic), . . . the same objects, which had been kept carefully from him were again presented to his notice. He could at once perceive a difference in their shapes; though he could not in the least say which was the cube and which the sphere, he saw they were not of the same figure. It was not until they had many times been placed in his hands that he learnt to distinguish by the eye the one which he had just had in his hands, from the one placed beside it. He gradually became more correct in his perception, but it was only after several days that he could or would tell by the eyes alone, which was the sphere and which the cube. . . . Of distance he had not the least conception. He said everything touched his eyes, and walked most carefully about, with his hands held out before him, to prevent things hurting his eyes by touching them. Great care was required to prevent him from falling over objects, or walking against them. Improvement gradually went on, and his subsequent sight was, and now is, comparatively perfect (1, 446–448).

Most persons would suppose that these boys attained to distance vision by merely focusing and using their eyes upon the objects of their environments. But this is far from the case. Their ocular experience was supplemented, corrected,

verified by manipular experience, and by their locomotor experience in getting their hands within reach of things. This part played by the hands in supplying the meaning of our retinal stimulations has probably never been better stated than by Professor Mead:

> What he (man) sees and hears means what he will or might handle. The proximate goal of all perception is what we can get our hands upon. If we traverse the distance that separates us from that which we see or hear and find nothing for the hand to manipulate, the experience is an illusion or a hallucination. The world of perceptual reality, the world of physical things, is the world of our contacts and our manipulations, and the experience of the eye and ear means first of all these physical things (4, 182).

But the problem is not yet solved, of how we become able, even with the aid of manipulation and locomotion, to translate retinal stimulations into a guarantee of external realities. A further contribution can be made by use of the interpretation that we have made of after-imagery.

We have noted that every case of after-imagery seen upon a ground involves two inverted cones or pyramids of light. Of these two physical figures, their bases are the only directly detectable or perceptible features; that of the inner figure, by its stimulation of the retina; that of the external figure, by touch and manipulation. To be sure, in cases of after-imagery we do not touch or manipulate the base of the figure as such, but only the surface upon which it rests, or from which the light proceeds to activate the retinae. But in cases of original vision, as of the object from which an AI is produced, the case is different. That object is itself, as such, tangible and more or less manipulable.

Here, then, we have a further cue to an understanding of distance vision. It consists in the fact that we find and substitute the bases of our external cones of light—or the objects from which those cones proceed—for the homologously-shaped but inverted bases of their correspondent inner figures. We learn by experience the correspondence between the bases of the two light-figures thus involved, and we learn to substitute the base of the external, for that of the internal figure. It is by a learned correlation of manipular experience with the small affected patches of our retinae that these latter take on

that appearance of definite external magnitude, which the boy of our first surgical report had not yet learned to experience. And while the brightness and color aspects of our visual experience are matters of retinal commotion—this, surely, rather than of synaptic functions in the occipital cortex, yet we learn to refer these to the apposite inverse positions in the base of our external light-figures, which they mean or point to. That is why our external objects themselves appear to be variously bright and colored. Such, I suggest, is an apposite description of the process of external or distance vision—a physically structured process throughout.

Now we can understand why our visual objects do not appear to us as being inverted, as the retinal patches affected by light from our external 'events' are. We substitute those events themselves, as the bases of our external light-figures, for the bases of the correspondent though inverted internal light-figures; and, through the mediation of the two light-structures, we react to those substitutes, rather than to the retinal stimulation as such.[1] The young infant starts off upon his way to adequate visual experience by exploration, touching, and manipulation of the objects of his near environment. Here, the bases of his external and internal light-figures are so large and so correlative in magnitude as to afford him adequate cues for differentiating among his objects, even when he does not get his hands upon them. Thence gradually he learns to interpret the more and more minute patches of retinal stimulation from more and more distant objects, by motor-imaginal substitution of the bases of his external light-figures. Here he experiences secondary helps, in the varying degrees of muscular tension in focusing his eyes and postural apparatus, and in the relative amount of haze intervening between his eyes and his far-off objects. But the basis of his distance vision is still describable as being a substitution of his external

[1] Some one may at first prefer to say that, having learned by the hands what is the meaning of our visual patches, we now substitute these present data for those meanings. But this, being largely the historical mode of stating the case, would perpetuate the historically false problem of why we accept an inverted datum for its obverse original. Quite obviously, the substitution is the other way around, as I have stated it.

for his internal light-figures. I am now putting it in this way, because in substituting the base of one for the base of the other he is in effect substituting the entire external figure for the internal one. We must state the case in this way, in order to appreciate the spatio-temporal or physical continuity of the process.

Now, too, we can explain why after-imagery seen without a ground is upright rather than inverted, whereas our only direct cue to interpretation is the inversely-formed patch of affected retinae. This is the effect of our long-established and responsible habit of interpreting such patches by reference to the upright objects from which the activating light came. This is the way we interpreted the form and position of the object from which the light has just come, and so we now interpret the persisting effect of that light, even when no light at all is affecting the retina. We are habituated to finding upright and external objects answering to our light-stimulated retinal patches, and the habit sticks.

That an AI seen with the eyes closed or blindfolded seems so noticeably smaller than its original is probably to be referred to the fact that there is no break or internal configuration of the now only apparent ground, from which its size can be judged. In the case of establishing after-imagery in a dark room, there is no such break or configuration of ground surrounding the original object; hence there seems to be less discrepancy in the size of the AI here resulting, than in cases where fixation took place in a lighted room, with a visible frame of reference, and the eyes then closed or blindfolded. Still, when in a dark room, AIs seem to be a bit smaller than the stimulus object. Perhaps a part of the explanation must be sought in reference to the fact that, for effective after-imagery, we must have our eyes closer to the object than we are accustomed to have such objects, in normal vision. Possibly, too, the ocular media so bend the rays of light that the basal angles of the inner cone are greater than those of the external cone. In such case, the retinal patch, or image, would be proportionally smaller than the base of the external cone; and then the bare AI, 'seen' without a ground, would

naturally appear smaller than the stimulus object. Yet this would not account for the apparent discrepancy of size in AIs without a ground, 'seen' in a lighted and in a dark room. This, however, was not precisely a part of my problem.

Summary

This study and account of after-imagery was prompted by a notion that there was a higher degree of continuity between distance vision and after-imagery than has been generally understood. The particular point of departure was a reaction against a notion that after-imagery did not exhibit a determinate perspective. The study took further color as being a reaction against the use of the word 'projection' as a supposedly apposite term for describing the processes of both after-imagery and distance vision.

For the purpose of securing data for substantiating these objections, a varied experimental study of after-imagery has been made and here briefly reported. Here it has been found and shown that after-imagery, 'seen' upon a ground, has a perspective of its own, inverse to that of original or distance vision. Both types of perspective have been interpreted in physical-geometrical terms, rather than in terms of the traditional 'pure' Euclidean geometry, and of Newtonian mechanics, which involve the meaningless concept of 'action at a distance.' The results of the experiment were so definite and clear, and their interpretation so illuminative of the nature of distance vision, that an interpretation of this process has been made the main theme of the paper.

Interpretation of the distance aspect of vision has generally been attempted, in the past, on the assumption that it must be done by reference to retinal stimulation alone, as if the eye in some mysterious way 'projected' its sensations or images upon its objects or grounds. In this way, a disturbing but false problem has been formulated, of why our objects, as seen, appear in an upright, obverse position. Contrary to this interpretation, it has been held here that such vision involves an active, physical continuum intervening between the object and the agent, and that both ends or bases of this

continuum must be taken into account, as physical things. Through this continuum, a two-way reference takes place. On the one hand, light from the object passes to the retinae and activates them—a commonplace of our thinking about vision. On the other, a locomotor, projicient, manipular, tactual reference is made by the agent, towards the object. In this latter process, the agent substitutes the object, as the base of an external cone or pyramid of light, for the patch of retinal affection, as the base of an inverse but homologously-shaped, inner light-figure. Each detail of the retinal image, or its retinal effect, he refers to the inverse but correlative detail of the object; and thus his object is upright, and he innocently *refers* his sensations to it—a very different matter from 'projecting' them upon it. If any one still supposes that we somehow 'project' a non-physical content upon a physical ground and then react to it, let him explain, in terms acceptable to any science, how we can react to anything that is non-physical.

(Manuscript received April 9, 1934)

References

1. Berkeley, George, Works, collected and edited by Alexander Campbell Fraser, in three volumes: Vol. I, pp. 27–112, 371–400, 444–448. (The report of the second surgical case is taken from Thomas Nunneley's *On the Organs of Vision: Their Anatomy and Physiology*, London, 1858, pp. 31–33.)
2. Allport, G. W., Eidetic imagery, *Brit. Jour. Psych.*, **15** (1924–25), 99–120.
3. Emmert, S., Grössenverhältnisse der Nachbilder, *Klinische Monatsbl. f. Augenheilkunde*, **19** (1881), 443–450; also quoted in *Genet. Psychol. Monogr.*, **1** (1926), 186.
4. Mead, George Herbert, The nature of the æsthetic experience, *Int. Jour. Ethics*, **36** (July, 1926), 382–393.

A CORRELATION FORM WITH ABSOLUTE CHECKS

BY ROBERT C. TRYON

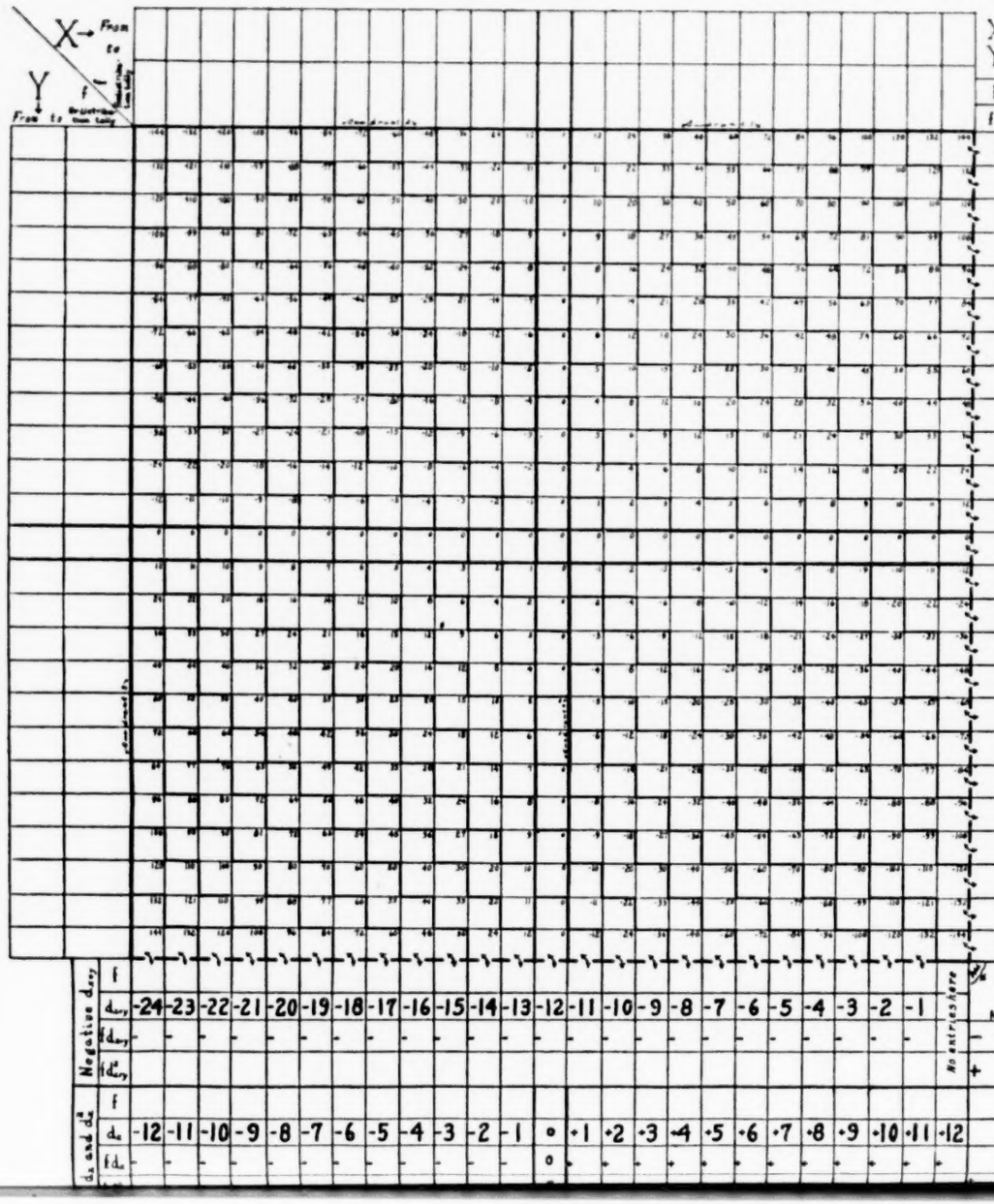
University of California

For several years, the writer has been addicted to the use of one or another of the innumerable correlation forms on the market. But with the use of every such form, he inevitably came to nurse a grudge against its author, either because of some undesirable feature present or of some desirable feature absent. From this cumulating affective and cognitive welter gradually became abstracted a concept of an 'ideal correlation form,' and the irrepressible consequence has been the production of one more correlation form to be added to the host already available. Figure 1 shows the front or grid side of this new form, Fig. 2, the reverse side.

Below are given what the writer considered the more important necessary properties of an 'ideal' correlation form, and following each such ideal feature is presented a statement of how the new form attempts to embody it.

1. A checking system should be provided, and one which will enable the worker to check the original plotting of the X, Y entries as well as the calculations leading up to and including r. Such checks should be absolute in the sense that all values should be calculated in two *independent* ways. The common practice of working over one's calculations twice does not constitute an adequate check. The writer has found, in his own case, that if on the first calculation some perverse mental imp impells him to believe that 9 times 8 is 73, then on recalculation the probability is great that the same devious impulse will lead him to 73 again. He mistrusts his own recomputations as a check, and, frankly, those of other people as well. An equational check is the only adequate check.

In the New Form, a Rigorous Checking System is Provided for All the Work.—The plot of X, Y scores is checked by means of the inclusion of an extra Y column next to the Y scale; after



X → From ... to ...

Y ↓ From ... to ...

X is _____

Y is _____

Positive d_{x+y}			d_y and d_y^2			fd_{xy}	
f	d_{x+y}	fd_{x+y}	f	d_y	fd_y	fd_y^2	+	−
	24			12				
	23			11				
	22			10				
	21			9				
	20			8				
	19			7				
	18			6				
	17			5				
	16			4				
	15			3				
	14			2				
	13			1				
	12			0	0	0	0	0
	11			-1				
	10			-2				
	9			-3				
	8			-4				
	7			-5				
	6			-6				
	5			-7				
	4			-8				
	3			-9				
	2			-10				
	1			-11				
0 0 0				-12				

Negative d_{x+y}: f | d_{x+y}: -24 -23 -22 -21 -20 -19 -18 -17 -16 -15 -14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 | fd_{x+y} | fd_{x+y}^2

d_x and d_x^2: f | d_x: -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 +1 +2 +3 +4 +5 +6 +7 +8 +9 +10 +11 +12 | fd_x: - - - - - - - - - - - - 0 + + + + + + + + + + + +

No entries here

+ + + | N | Σd_y | Σd_y^2 | + − | Σd_{x+y}

N | M_x | σ_x | − | + | Σd_{x-y} | σ_y | M_y

Σd_x | Σd_{x-y}^2

FIG. 1. The front or grid side of the correlation form.

CORRELATION FORM

WITH ABSOLUTE CHECKS OF ALL OPERATIONS INCLUDING PLOTTING

For order of operations and supplementary formulæ, see reverse side of this form

Work straight from top to bottom

Check 1: $\Sigma d_x + \Sigma d_y =$ $= \Sigma d_{x+y} =$

Check 2: $\Sigma d_x^2 + \Sigma d_y^2 + 2\Sigma d_{xy} =$

$= \Sigma d_{x+y}^2 =$

$\Sigma d_x/N = c_x =$; $c_x^2 =$

$\Sigma d_y/N = c_y =$; $c_y^2 =$

$\Sigma d_{x+y}/N = c_{x+y} =$; $c_{x+y}^2 =$

Check 3: $c_x + c_y = c_{x+y}$; $c_x^2 + c_y^2 + 2c_xc_y = c_{x+y}^2 =$

$\Sigma d_x^2/N =$; $\Sigma d_{xy}/N =$

$\Sigma d_y^2/N =$; $\Sigma d_{x+y}^2/N =$

Check 4: $\Sigma d_x^2/N + \Sigma d_y^2/N + 2\Sigma d_{xy}/N =$ $= \Sigma d_{x+y}^2/N$

Arbitrary origins: $A_x =$; $A_y =$

Class intervals: $I_x =$; $I_y =$

$M_x = A_x + (I_xc_x) =$

$M_y = A_y + (I_yc_y) =$

Check 5: $M_x + M_y =$

$= A_x + A_y + (I_xc_x) + (I_yc_y) =$

$\sigma_x^2 = I_x^2(\Sigma d_x^2/N - c_x^2) =$; $\sigma_x =$

$\sigma_y^2 = I_y^2(\Sigma d_y^2/N - c_y^2) =$; $\sigma_y =$

Let $E = \sigma_x^2/I_x^2 + \sigma_y^2/I_y^2 =$

Let $F = \Sigma d_{x+y}^2/N - c_{x+y}^2 =$

Let $G = \Sigma d_{xy}/N - c_xc_y =$

Check 6: $E = F - 2G =$

$r_{xy} = (1/\sigma_x\sigma_y)(I_xI_y[(F-E)/2]) =$

$r_{xy} = (1/\sigma_x\sigma_y)(I_xI_yG) =$

Check 7: The two above values for r_{xy} should be equal.

Correction (Sheppard's) for broad categories (underlined):

$\varrho_x^2 = \sigma_x^2 - .0833I_x^2 =$; $\varrho_x =$

$\varrho_y^2 = \sigma_y^2 - .0833I_y^2 =$; $\varrho_y =$

$\varrho_{xy} = r_{xy}\sigma_x\sigma_y/\varrho_x\varrho_y =$

Composite constants: $M_{x+y} = M_x + M_y =$

$\varrho_{x+y} = \sqrt{\varrho_x^2 + \varrho_y^2 + 2\varrho_{xy}\varrho_x\varrho_y} =$

Probable errors: Let $K = .6745/\sqrt{N} =$

$P.E._{M_x} = K\varrho_x =$; $P.E._{\sigma_x} = .7071 P.E._{M_x} =$

$P.E._{M_y} = K\varrho_y =$; $P.E._{\sigma_y} = .7071 P.E._{M_y} =$

$P.E._{M_{x+y}} = K\varrho_{x+y} =$; $P.E._{\sigma_{x+y}} = .7071 P.E._{M_{x+y}} =$

$P.E._{r_{xy}} = K(1-r_{xy}^2) =$

Order of Operations

It will pay to work carefully, for the calculations will then check and no further recomputation will be necessary.

Supplementary Formulae

Coefficient of alienation: $k_{xy} = \sqrt{1 - r_{xy}^2}$

Order of Operations

It will pay to work carefully, for the calculations will then check and no further recomputation will be necessary.

1. Determine the interval values, I_x and I_y.
   Process: Find the range of X scores; divide this by 25. If the quotient is an integer, this is I_x; if not an integer, raise quotient to next highest integer, and this is I_x. (If original X values are not integers, disregard decimal points and consider as integers, and go through same process above putting in proper decimal point in I_x.) Determine I_y in the same manner.
2. Enter class values under the "Y" and to the right of "X", being sure to have Y magnitudes go from low to high from the bottom to top, and X magnitudes go from low to high from left to right.
3. Plot the X,Y scores, using the more preferred tally system, viz., $\mathcal{NL}$ (=8). The plotting is facilitated by the use of an ordinary ruler held perpendicular to the Y scale, moving this up and down to locate the Y scores, then the slipping of a square card along the ruler (so that the left margin of the card will reach well up to the X scale) enables one to locate the corresponding X score at the top. In the cell at the intersection of the card and ruler the entry is made.
4. Redistribute the X scores alone in the frequency table labelled "X redistribution tally" Do the same with the Y scores. These redistributions serve only to check the plotting, as indicated in the next step.
5. Add the frequencies of each row and enter total in correct cell under column "f"(6th column from extreme right); then add frequencies of each column and enter total in correct cell in row "f"(4th row from the bottom). There are now two independently derived frequency distributions of Y, one under "Y redistribution tally", the other under the "f" column; and analagously there are two of X. If these distributions check, *then the plotting is checked*, but if discrepancies occur, then it is necessary only to check those X,Y rows and columns in error.
6. Add the frequencies of each of the upper-left-to-lower-right diagonals as follows:
   (a) First, do the *longest diagonal* and then the diagonals *above it*. Let the heavy lines guide the eye in adding up the frequencies of a given diagonal. Enter the total frequency of each diagonal in its proper cell under column "f" *next to the grid*, the double-headed arrows indicating the proper cell for a given diagonal.
   (b) Second, do the diagonals *below* the longest, entering the frequencies in the proper cells in the "f" row *just below* the grid, being guided to the proper cells by the double-headed arrows.
7. Note the number in the upper right corner of each cell in the grid. This is the d_{xy} product for that cell. In each row of quadrant 1, multiply the frequency in each cell times the number in the upper corner of the cell, and sum all these for all cells in that row, placing the total in the proper cell in the "+" column of the final set of columns headed by "fd_{xy}". Do the same for the rows of quadrant 2, placing the sums for each row in the proper cells in the "−" column. Similarly, the sums for quadrant 3 go in the "+" column, and those for quadrant 4 go in the "−" column.
8. Perform the other operations indicated in the columns to the right and beneath the grid; namely, in the proper cells compute fd_{xy}, and fd^2_{xy} (columns and rows), fd_y and fd^2_y, and fd_x and fd^2_x.
9. Sum these columns and rows as indicated, being sure to regard algebraic sign in all total cells except those involving the d''s which are always positive.
10. Perform the double sums as indicated in the three extra large cells at the lower right corner of the chart, being sure to regard algebraic signs in the cells marked Σd_{xy}, and Σd_{xy}.
11. Proceed now with the calculations indicated at the right of the chart, working straight from the top to the bottom, and calculating the checks in their proper places. *Each check formula checks all preceding operations.* Whenever a check fails, it will only be necessary to work back to the last check.

FIG. 2. The reverse side of the correlation form.

Supplementary Formulae

Coefficient of alienation: $k_{xy} = \sqrt{1 - r^2_{xy}}$

Prediction formulae:

$$\bar{X} = r_{xy}(\sigma_x/\sigma_y)(Y-M_y) + M_x \qquad \bar{x} = r_{xy}(\sigma_x/\sigma_y)y$$

$$\bar{Y} = r_{xy}(\sigma_y/\sigma_x)(X-M_x) + M_y \qquad \bar{y} = r_{xy}(\sigma_y/\sigma_x)x$$

$$\sigma_{x.y} = \sigma_x k_{xy} \qquad \sigma_{y.x} = \sigma_y k_{xy}$$

Correlation ratio:

$$\eta_{xy} = \sigma_{\text{means of rows}}/\sigma_x \qquad \eta_{yx} = \sigma_{\text{means of columns}}/\sigma_y$$

$$P.E._{\eta} = (.6745/\sqrt{N})(1-\eta^2)$$

Test for linearity: η^2-r^2; $P.E._{\eta^2-r^2} = (.6745/\sqrt{N})(2\sqrt{\eta^2-r^2})$

Probable error of differences:

$$P.E._{M_x-M_y} = \sqrt{(P.E._{M_x})^2 + (P.E._{M_y})^2 - 2r_{xy}(P.E._{M_x})(P.E._{M_y})}$$

$$P.E._{\sigma_x-\sigma_y} = \sqrt{(P.E._{\sigma_x})^2 + (P.E._{\sigma_y})^2 - 2r^2_{xy}(P.E._{\sigma_x})(P.E._{\sigma_y})}$$

Determination of the reliability of individual measurement:

(a) If X and Y are random ("split") halves of the total (X+Y) score:
Reliability coefficient of (X+Y) = $r_{x+y} = 2r_{xy}/(1+r_{xy})$
$P.E._{r_{x+y}} = (.6745/\sqrt{N})[2(1-r_{x+y})]$

(b) If X and Y are two independent measures of a variable, such as two forms of a test:
Reliability coefficient of X or Y = r_{xy}; $P.E._{r_{xy}} = (.6745/\sqrt{N})(1-r^2_{xy})$

(c) Estimated reliability coefficient of a variable which consists of n times the material of the measurement which has a reliability coefficient, r:
Estimated reliability coefficient = $r_{nn} = nr/[1+(n-1)r]$
$P.E._{r_{nn}} = (.6745/\sqrt{N})\,[n(1-r^2)]/[1+(n-1)r]^2$

(d) P.E. of an individual's score in a measurement which has a reliability coefficient, r:
$P.E._{\text{individual score}} = .6745\sigma_{\text{distribution}}\sqrt{1-r}$

(e) Index of reliability: $\sqrt{r}$

(f) Sigma of "true scores": $\sigma_{\infty} = \sigma_{\text{distribution}}\sqrt{r}$

Correction for attenuation:
Where the reliability coefficients of X and Y are r_x and r_y, respectively,
$r_{\infty_x\infty_y} = r_{xy}/\sqrt{r_x r_y}$; $P.E._{r_{\infty_x\infty_y}} = (.6745/\sqrt{N})r_{\infty_x\infty_y}\sqrt{r^2_{\infty_x\infty_y} + 1/r^2_x + A_x + A_y}$
For the values, A_x, A_y, see T.L.Kelley, *Statistical Method*, formula [101] and Table XL.

Correction for range:

(a) Of the reliability coefficient:
The reliability coefficient, r_x, is calculated for the range indicated by σ_x. For the new range with standard deviation, Σ_x, and letting $\Sigma_x/\sigma_x = h$, the new reliability coefficient, R_x, will be
$R_x = 1 - (1-r_x)/h^2$

(b) Of the correlation coefficient:
The correlation coefficient, r_{xy}, is calculated for the range indicated by σ_x. For the new range with standard deviation, Σ_x, and letting $\Sigma_x/\sigma_x = h$, the new correlation coefficient, R_{xy}, will be
$R_{xy} = r_{xy}h/\sqrt{(1-r^2_{xy}) + (r_{xy}h)^2}$
$P.E._{R_{xy}} = .6745\sigma_{r_{xy}}[R_{xy}(1-R^2_{xy})/r_{xy}(1-r^2_{xy})]$

the plot is made, in this extra column the Y scores are redistributed without reference to the X scores. If the frequencies in this extra frequency table agree with those in the grid-derived 'f' column from which the Y mean and sigma are calculated, and if the two analogous frequency tables of the X scores also agree, *then the plotting is checked.* By equations, all the later calculations leading to r are checked as one proceeds with the calculations. Thus, if the first check equation is satisfied, then all operations previous to it are correct. Calculations are then made up to the next check equation, which, if satisfied, indicates that all operations to this new point are correct. Consequently, whenever a check fails, it is necessary only to work back to the last check. If all check equations are satisfied, *then no recomputations are necessary.* A careful worker thus never has to retrace ground, and at the end he knows his work is errorless. The check equations are possible because r is determined independently by the product and sum methods (using deviations from arbitrary origins), these methods being worked simultaneously. The values used in calculating the means and sigmas are secured from the usual f, d, fd, and fd^2 tables. The products, $d_x d_y$, for which the notation d_{xy} is used, are printed in the upper right corner of cells of the grid in the usual way, and the customary '+' and '−' columns for sums of d_{xy} are given at the extreme right. The sum $(d_x + d_y)$, for which the notation d_{x+y} is used, is the same in each upper-left-to-lower-right diagonal, and cells adjacent to the right and below the grid are provided to receive the frequency of each diagonal. After summing the diagonals, the operator has only to calculate the fd_{x+y} and $fd^2{}_{x+y}$ as indicated in Fig. 1. After these operations and the grand total of the columns and rows have been made, the order of operations are executed as shown at the extreme right of the form. Checks 1 and 2 check the accuracy of the grand totals and the operations leading to them. Check 3 checks the corrections, c_x, c_y, check 4 the divisions by N, check 5 the means, check 6 the sigmas, and check 7 the correlation coefficient, which is calculated from the products, d_{xy}, and from the squares of the sums, $d^2{}_{x+y}$.

2. The number of rows (or columns) of the grid should be at least 25. Some correlation forms possess not more than 18 rows (or columns), a feature which makes such forms useless to a person working with scores the range of which happens to be 19 or 20, for then he must use an interval value, I, of 2, and he finds that he can utilize only 10 categories. If the scores tend to slight skewness, or if a rare widely divergent score occurs, most of the scores would thus fall in only four or five categories.

In the New Form the Grid Contains 25 Rows and 25 Columns.—Thus if the range of the X (or Y) scores is 26, then an I of 2 will still leave 13 categories.

3. Accessory equipment should not be required, such as stencils or guides for adding diagonals. Such devices generally get mislaid or damaged. The correlation form should be complete in itself.

The new form requires no accessory device to calculate the sum of the diagonals, for either the upper or lower margin of each diagonal is indicated by an extra heavy line. With a little practice in trailing a pencil along this margin, one can sum up without difficulty the frequencies of a given diagonal.[1]

4. Cells in the grid should be large enough for at least 15 entries.

The New Form Satisfies this Requirement.—If one works with a very large N and requires more cell space, he can fractionate his population into sub-samples of, say, 200 cases each, use a single form for each such sub-sample, apply checks 1 and 2 to each, and then enter the checked grand totals into a master form on which the final calculations can be made.

5. In the formulæ from which the means, sigmas and r are calculated, only standard statistical symbols should be employed. The common technique in many forms of letting A, B, C, D, etc., stand for various sums introduces new symbols in an already highly symbolized system, and since each such new symbol is not descriptive of the value which it represents, possibility of confusion and error arises. An operator

[1] In the first reprinting of this correlation form the cells of every other diagonal will be marked off distinctly by visual markers, thus greatly facilitating the summing up of a given diagonal frequency.

rarely knows what operations he is performing when he manipulates these new symbols, and such 'blind' work is not conducive to intelligent accurate work.

Standard Symbols are Used Throughout in the New Form Except in a Few of the Last Calculations Resulting in r.—Here the symbols are employed only to avoid reprinting several long expressions that are given immediately above.

6. Space and formulæ should be provided for probable errors of all constants, and for corrections of sigmas and r for broad categories.

The new form provides these features in the space below check 7.

7. A complete description of the order of operations should be provided, so that an operator can always ascertain whether or not he is proceeding correctly. Furthermore, such complete directions will enable any intelligent neophyte to work 'on his own hook' without wasting the time of the local statistical expert in initiating him into the mysteries of the form. Such printed directions greatly reduce the time required in teaching the use of a form in an elementary statistics class.

On the reverse side of the form, as indicated in Fig. 2, a complete order of operations is provided.

8. For ready reference, a valuable adjunct would be the presence on the form of important formulæ pertaining to zero order r, such as formulæ for prediction, corrections for length of the test, for attenuation,[2] for range, *with their probable errors.*

On the Reverse Side of the New Form (Fig. 2) *are Given These and Other Relevant Formulæ with Probable Errors.*—Where X and Y are 'split-halves' of a measurement, the mean and sigma of the total $(X + Y)$ measure is usually required. On the front side of the form under Sheppard's correction are given the formulæ for this mean and sigma. Incidentally, where the frequency distribution of the $(X + Y)$ total score is desired, this may be secured from the d_{x+y} frequency table.

9. To have general utility, a form should be inexpensive.

[2] For a recent consideration by T. L. Kelley of the error of an r corrected for attenuation see Bancroft Beatley, *Achievement in Junior High School* (Harvard University Press), p. 43.

The cut from which the form is printed was produced by a simple photo-process, hence the form is produced more cheaply than the ordinary form that is set up in type.[3]

10. The general size of the form should be of such proportions that it will fit in with other data on paper of standard size.

When folded, the new form has the standard dimensions of $8\frac{1}{2}$ *by* 11 *inches.*

11. The paper on which the form is printed should have a tint and surface which will reduce eye-strain, a frequent accompaniment of close work over cross-sectioned paper.

The new form is printed on stiff yellow ledger paper.

(Manuscript received December 27, 1932)

[3] The forms may be secured from the Student's Coöperative Store, University of California, Berkeley, Calif.